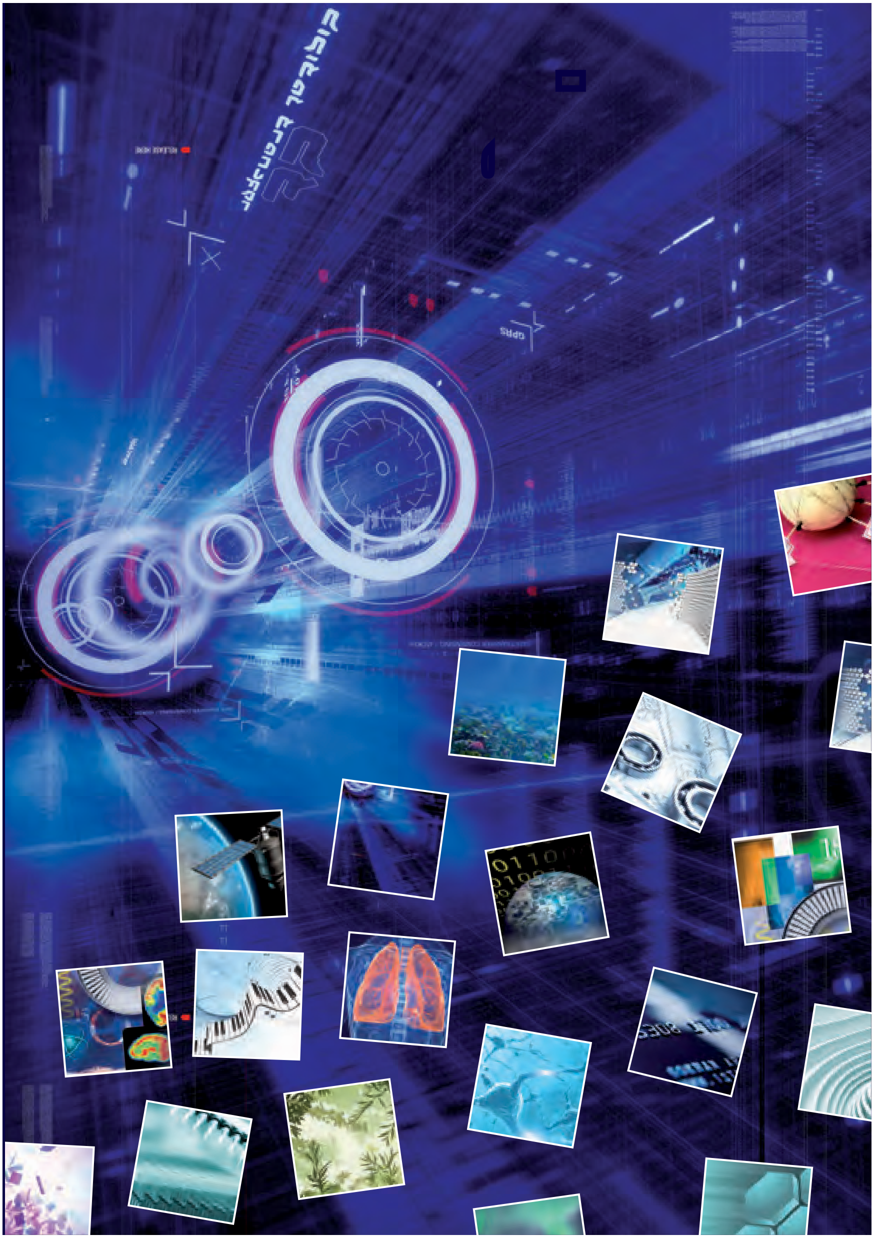
Matematica, l'esplosione continua



ISBN 978-88-96336-19-9

Stampa: Monograf S.r.l., Bologna

I matematici sono richiestissimi, la matematica è applicabile a qualsiasi cosa, un matematico trova un lavoro stabile nei tre anni dopo la laurea ecc..., tante belle parole che distruggono l'idea comune (ad es., nei miei tempi liceali) che un laureato in matematica possa fare soltanto l'insegnante. Hanno provato anche con il cinema e con le TV a far capire al mondo che i matematici possono essere utili anche al di fuori di una scuola.

I matematici trovano quasi subito lavoro, nel settore bancario, assicurativo, editoriale, di consulenza, perfino nella polizia scientifica oltre che, naturalmente, nell'insegnamento. I matematici si sentono obbligati a "continuare a studiare" anche dopo la laurea, non necessariamente ancora matematica (ma informatica, economia, biologia, ...), e cercare sempre di utilizzare quel modo di ragionare che li caratterizza in quello che studiano e fanno. Proprio qui sta il problema e la virtù di un matematico: cercare di adattarsi ai vari mestieri cercando di utilizzare al meglio quello che ha studiato. Per un ingegnere, un medico, un avvocato è molto più semplice, sanno già cosa faranno una volta finiti gli studi; il loro lavoro, il più delle volte, sarà strettamente legato ai loro studi, sfrutteranno necessariamente quello che hanno studiato, non avranno bisogno di adattarsi a realtà che apparentemente non gli appartengono. Per un matematico invece è tutto il contrario: perché, partendo dalla sua formazione e dalle cose che ha appreso, il matematico dovrà confrontarsi con problemi concreti (ed anche quelli dell'insegnamento lo sono!) e cercare di risolverli adattando le proprie conoscenze alla realtà: quella del matematico è quindi la scelta più coraggiosa ed esaltante e va ricordato il fatto che una laurea in matematica richiede probabilmente anche più impegno di una qualsiasi altra laurea.

Oltre a teoremi, sempre più ricercatori matematici si applicano alla ricerca di algoritmi, cioè escogitano e mettono a punto procedimenti di calcolo: lo sviluppo dei computer stimola – anziché limitare – il lavoro di molti matematici, rendendo realizzabili modelli che decenni fa sarebbero stati impraticabili. Perciò molti matematici *lavorano in enti di ricerca (per esempio* Cnr, Enea, Esa, ...*) in diversi ambiti scientifici che richiedono modellistica matematica. È qui che si manifesta maggiormente l'importanza della continuità dello spettro: da matematica "pura" a matematica ispirata dalle applicazioni ad applicazioni concrete.*

È totalmente falso che iscrivendosi a Matematica si limiti il proprio futuro all'insegnamento o poco più. Alcuni faranno "ricerca" matematica, altri come detto, si dedicheranno all'insegnamento ovvero alla cura dei figli dei nostri concittadini, ed altri al vasto campo delle applicazioni (ed in tal caso sarebbe importante avere nel piano degli studi, al di la' dei corsi di matematica, anche corsi di informatica e di statistica: fra i corsi a scelta sono più utili questi che i corsi di chitarra messicana o insegnamenti con titoli fantasiosi).

Oltre a diventar matematici è importante che i matematici riescano a farsi conoscere, come tali ovvero come i migliori "affrontatori" di problemi di qualsivoglia natura in campo scientifico. Le Società di matematica di Francia, allo scopo di evidenziare questo fatto in molte sfaccettature delle attività più importanti della nostra vita quotidiana, hanno pubblicato nel 2002 un volumetto intitolato "L'explosion de Mathématiques" del quale l'Unione Matematica Italiana ha ottenuto a suo tempo i diritti di traduzione e di pubblicazione. Nel 2013 è uscito il "seguito" ovvero un volumetto intitolato "Mathématiques: l'explosion continue" l'Unione Matematica Italiana e la Società Italiana di Matematica Applicata e Industriale hanno chiesto, ottenendoli, i diritti di pubblicazione e, grazie al collega Marco Maggesi, che ha curato la traduzione (e che qui sentitamente ringraziamo), siamo in grado di presentare questo nuovo volumetto che, ne siamo convinti, renderà interessante la lettura da parte dei giovani interessati a coltivare la nostra disciplina e fornirà motivi di positiva sorpresa a chi, pur addetto ai lavori, non conosceva nei dettagli il meraviglioso mondo delle possibili applicazioni della matematica.

Giuseppe Anichini

Si ringraziano
Société Mathématique de France
Société de Mathématiques Applicquées et Industrielles
Société Française de Statistique

per averci concesso i diritti di traduzione, di riproduzione e diffusione gratuita.

Sommario

Premessa

Maria J. Esteban, *presidente della Société de Mathématiques Appliquées et Industrielles 2009-2012*
Bernard Helffer, *presidente della Société Mathématique de France 2010-2012*
Jean-Michel Poggi, *presidente della Société Française de Statistique 2011-2013*

Da una decina d'anni, numerose iniziative sono apparse in Francia per capire meglio il ruolo della matematica nella nostra società. I matematici sono diventati più consci del fatto che era necessario far conoscere meglio la specificità e l'utilità della loro disciplina. Una delle prime iniziative è stata la pubblicazione di Explosion des Mathématiques *nel 2002. Questa raccolta è stata ampiamente diffusa e tradotta in diverse lingue*[1]*. È ormai esaurita, anche se è ancora disponibile sul web*[2]*. I circa dieci anni che ci separano dalla prima edizione hanno visto un rapido sviluppo di tutti i rami della matematica e il loro sviluppo in crescita in tutti i settori della società: l'esplosione continua! Il congresso* Maths à Venir, *il cui scopo era quello di testimoniare l'espansione crescente della matematica, è stato nel 2009 un vero successo, al quale hanno aderito molti soggetti del mondo dell'industria. I fondi raccolti in questa occasione ci permettono di realizzare una nuova edizione. È per questo che abbiamo deciso di pubblicare un nuovo opuscolo che riprende tre articoli aggiornati dalla vecchia brochure e ventuno*[3] *nuovi articoli che tengono conto dei recenti sviluppi e della diversità crescente delle interazioni e delle applicazioni della matematica.*

---

[1] Traduzione italiana, a cura dell'UMI: http://umi.dm.unibo.it/downloads/explosio.pdf
[2] Versione originale in francese: http://smf.emath.fr/Publications/ExplosionDesMathematiques/
[3] Il numero si riferisce all'edizione originale. Nella traduzione presente non sono stati considerati due articoli relativi alla matematica francese.

Prima di lasciare che il lettore scopra questi articoli, collochiamo questo volume nel suo contesto. Viviamo ancora oggi una situazione paradossale. La matematica è uno strumento insostituibile di formazione al rigore e al ragionamento; sviluppa l'intuizione, l'immaginazione, il pensiero critico, oltre ad essere un linguaggio universale e costituire un elemento forte della cultura. Inoltre svolge, per le sue interazioni con le altre scienze e per la sua capacità di descrivere e spiegare fenomeni complessi messi in evidenza nella natura e nel mondo tecnologico, un ruolo crescente nella nostra vita quotidiana. Questo fatto è spesso ignorato o almeno poco considerato dalla maggioranza delle persone, per le quali la matematica è una disciplina astratta, che non si evolve, congelata in una prospettiva di apprendimento e che ha poco a che fare con il mondo reale.

Naturalmente, si possono trovare a questa contraddizione spiegazioni che tengono conto della natura specifica della matematica. Si tratta di una disciplina che oltre ad alimentarsi attraverso i legami con le altre scienze, con la società e con il mondo dell'industria, si auto alimenta: nuove teorie non distruggono le precedenti, ma le utilizzano, le migliorano o le superano. Anche se molti ricercatori in matematica sono principalmente interessati al lato intellettuale o persino estetico della disciplina, ce ne sono molti altri che sono coinvolti in altre direzioni di ricerca legate alle sue reali applicazioni. Così le applicazioni arricchiscono la ricerca in matematica, ma da sole non possono guidarla. Questo sottile equilibrio tra i fattori interni ed esterni del suo sviluppo deve essere assolutamente preservato. Voler definire o misurare l'attività o la ricerca in matematica attraverso le applicazioni esistenti o potenziali ne provocherebbe la loro scomparsa. Viceversa, privilegiare l'assiomatizzazione, lo studio delle strutture e delle dinamiche interne della disciplina, come ha fatto, certamente con grande successo, la matematica francese nel periodo 1940-1970, ha portato in Francia a ritardare lo sviluppo della matematica detta applicata, diversamente da quello che stava accadendo nello stesso periodo negli Stati Uniti e in Unione Sovietica. I fattori di progresso si trovano molto spesso alle frontiere della disciplina. Oggi, e ne siamo lieti, la matematica ha ristabilito legami forti con le scienze economiche e con le altre scienze, e, talvolta, ne ha creati di nuovi. Il confine tra matematica pura e matematica applicata è diventato permeabile: la matematica fondamentale viene impiegata per risolvere problemi concreti sempre più difficili, mentre nuovi problemi teorici vengono formulati a partire dalle applicazioni.

Lo scopo di questa pubblicazione è quello di far scoprire tutte le attrattive ed i vantaggi del mondo matematico e, soprattutto, la notevole efficacia della *matematica nella soluzione di problemi sociali e tecnologici; di mostrare la matematica nei sui aspetti più diversi – scientifico, tecnico, culturale, sociologico; evidenziare la diversità e l'universalità di una disciplina che ha legami sia con la fisica, la chimica, la meccanica, l'informatica, l'economia e la biologia, che con le discipline umanistiche e le scienze sociali, attraversando tutti gli aspetti della tecnologia. La matematica è ovunque. Senza di lei non ci sarebbe nessun computer, nessun sistema di informazione, nessun telefono cellulare, nessun laboratorio di progettazione per i costruttori di automobili e per i sistemi aerospaziali di tracciamento; e così via per l'elaborazione dei segnali satellitari, la decodifica del genoma, le previsioni meteorologiche, la crittografia, le smart card, i robot, i motori di ricerca e l'elaborazione di grandi quantità di dati ...*

Al di là del suo ruolo nella scienza accademica e nella formazione di base, la matematica è onnipresente nella scuola e nella società di oggi. Essa segue, accompagna e qualche volta precede, lo sviluppo scientifico e tecnologico attuale, il quale fa riferimento tanto ai risultati di ricerca fondamentale contemporanea più recente, quanto alle scoperte accumulate nel passato.

Infine, le necessità nella matematica crescono con l'accelerazione e la modificazione delle creazioni tecnologiche. È impossibile evitarlo, dato che siamo di fronte alla necessità di sviluppare, gestire, analizzare e ottimizzare sistemi sempre più complessi. Gli Stati Uniti lo hanno capito bene, dato che l'NSF (National Science Foundation, l'organizzazione federale nordamericana responsabile della distribuzione dei fondi per la ricerca universitaria) ha notevolmente aumentato il suo sostegno finanziario alla matematica dal 2000. Più di recente, altri paesi emergenti come la Cina hanno investito massicciamente in questo settore.

È in questo contesto che ci sembra prezioso completare e aggiornare questo strumento di diffusione e di divulgazione della matematica. Auguriamo al lettore delle belle e stimolanti scoperte.

Tagliare, frollare, affettare, sfumare:

un racconto culinario sulla risoluzione informatica di problemi difficili

Nicolas Schabanel, *direttore di ricerca CNRS presso Université Paris Diderot*

Pierre Pansu, *professore presso Université Paris-Sud*

Come disporre gli ospiti in due tavoli senza mettere insieme coloro che si detestano? Cercare una soluzione a questo rompicapo che sia indipendente dal numero di ospiti è un problema estremamente difficile da risolvere quando il numero di ospiti cresce. Con l'aiuto di un po' di programmazione geometrica, possiamo trovare una risposta che approssima la soluzione ottimale con una precisione stimabile.

Questa sera invito alcuni amici a cena. Purtroppo, non ho un tavolo abbastanza grande e devo disporre le persone in due tavoli. Come formare i gruppi?

Fermamente deciso a risolvere questo problema in modo razionale, prendo carta e penna e chiamo Maelys, sempre la prima ad essere informata sugli ultimi pettegolezzi, per sapere chi ha litigato con chi ultimamente. Tenendo il telefono con una mano, mi ritrovo a rappresentare ogni ospite da un piccolo cerchio (chiamato *nodo* in teoria dei grafi) e tracciare una linea (un *arco*, è più chic) tra ogni coppia di ospiti che, in questo momento, è meglio tenere separati. Mi rendo subito conto che la situazione è piuttosto complicata: c'è molta più animosità tra di loro di quanto pensassi; per esempio, nessun gruppo sembra delinearsi chiaramente sul mio disegno (chiamato anche *grafo*).

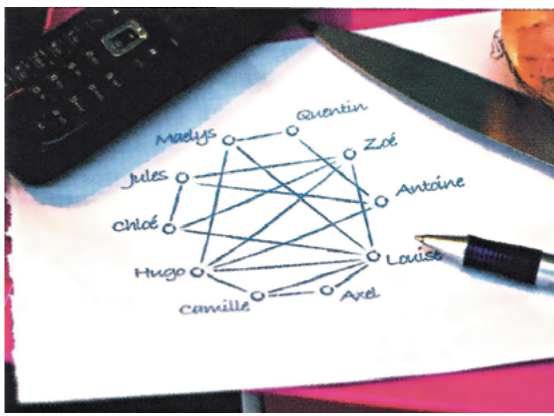


Figura 1. Il grafo delle incompatibilità tra i miei amici

Cerco, dunque, di trovare la suddivisione dei miei amici in due gruppi in modo che si possa ridurre al minimo il numero di inimicizie all'interno di ogni tavolo, il che equivale a massimizzare il numero di archi che vanno da un tavolo all'altro. Matematicamente, provo a *dividere* i vertici del mio grafo in due parti in modo da massimizzare il numero di archi tagliati dalla partizione, vale a dire, a cavallo tra i due tavoli. (Per non aggiungere complicazioni inutili, non cercheremo di ottenere due gruppi della stessa grandezza. D'altra parte, le tecniche di risoluzione con o senza questo vincolo non differiscono significativamente.)

Quando si mangia?

Se fossi un fisico, certamente costruirei un dispositivo di questo tipo: per ogni ospite prenderei una biglia, e per ogni arco porrei una grande molla compressa per simulare l'inimicizia tra le due biglie. Poi lancerei il tutto in un corridoio, le molle allontanerebbero le biglie che si detestano, le quali si disporrebbero, quindi, per tutta la lunghezza del corridoio e dunque otterrei la disposizione dei due tavoli dividendo il corridoio a metà: così avrei determinato la distribuzione degli invitati tra i due tavoli in funzione del punto del corridoio dove si è posata la loro biglia. Il problema è che, in questo modo, non disporrei di alcuna garanzia sulla qualità della soluzione trovata.

Un matematico potrebbe dimostrare molto facilmente che si può trovare una soluzione ottimale. In realtà, ho due scelte per ogni ospite: metterlo al primo o al secondo tavolo. Quindi non vi sono che $2 \times 2 \times \ldots \times 2 = 2^n$ modi possibili per suddividere n ospiti tra i due tavoli. È quindi sufficiente elencarli tutti e cercare quello che taglia il massimo numero di archi. Tuttavia, siamo sicuri che questo metodo sia migliore di quello precedente?

Certo, in questo modo si troverà la soluzione ottimale, ma dopo un tempo *esponenzialmente* lungo: 2^n tentativi. Ad esempio, con 60 ospiti, è necessario elencare e testare 2^{60} ossia circa 10^{18} partizioni, il che richiederebbe approssimativamente 10^9 secondi cioè circa 100 anni con un ipotetico processore capace di una frequenza di 100 GHz. E se avessimo invitato 61 persone, sarebbero 200 anni, e ci vorrebbero 400 anni per 62, ecc. Inutile dire che questa non è la soluzione adatta non appena il numero degli invitati cresce di poco: aggiungere un solo invitato raddoppia il tempo di calcolo (e questo *indipendentemente* dalla velocità del processore)! In altre parole, raddoppiare la velocità del processore non permetterebbe di risolvere il problema per il caso di un solo ospite in più!

Con 60 ospiti, sarebbe necessario elencare e testare 2^{60} ossia circa 10^{18} partizioni, il che richiederebbe approssimativamente 10^9 secondi cioè circa 100 anni con un processore ipotetico capace di una frequenza di 100 GHz.

Asciugare fino all'osso

È possibile trovare la migliore soluzione per il nostro problema in un tempo ragionevole? In qualità di informatico, risponderei che abbiamo buoni motivi per pensare che non è scon-

Metodi efficienti e inefficienti

In informatica, è opinione comune considerare un metodo di risoluzione *efficiente* se il numero di operazioni da eseguire è al *più polinomiale* in funzione della grandezza del problema (n^{α} nel caso peggiore per un problema di grandezza n, qualunque sia α). Per un metodo polinomiale, raddoppiare la grandezza del problema aumenta il tempo di risoluzione di un fattore costante $2^{\frac{1}{\alpha}}$: le variazioni del tempo di calcolo "seguono" quelle della grandezza dei dati in proporzioni limitate.

Al contrario, l'effetto delle grandezze su di un algoritmo di risoluzione in tempo esponenziale (per esempio 2^n) è catastrofico:

il tempo di calcolo di un algoritmo in tempo esponenziale raddoppia ogni volta che la grandezza del problema aumenta anche solamente di uno! ($2^{n+1} = 2 \times 2^n$). Ad esempio, se la potenza di calcolo dei computer raddoppia ogni 18 mesi, con un algoritmo polinomiale possiamo sperare di risolvere problemi $2^{\frac{1}{\alpha}}$ volte più grandi ogni 18 mesi, mentre con un algoritmo esponenziale, il massimo che si può sperare è di risolvere dei problemi che hanno una grandezza +1 ogni 18 mesi ...

Il numero di operazioni in funzione della grandezza n dell'input del problema è una misura della complessità *intrinseca* di un problema: *indipendentemente dalla macchina utilizzata* più questa funzione cresce rapidamente, più sarà difficile risolvere questo problema per grandi dimensioni.

tato perché questo problema (il cui nome scientifico è *Max-Cut*) è "uno dei più difficili". Ecco cosa intendo dire con questo.

Per risolvere un problema con un computer, deve essergli fornita una ricetta (*programma*) che lo risolva e i valori dei parametri (nel nostro caso: il numero di ospiti e l'elenco degli archi di inimicizia tra loro). Per scrivere questa ricetta (voglio dire, questo *programma*, noto anche come *algoritmo*), possiamo partire da zero, o sfruttare altre ricette per evitare di perdere troppo tempo (rischiare di introdurre nuovi bug). Ad esempio, si può immaginare che la ricetta che risolve il problema ipotetico *A* = "far uscire questo robot dalla stanza" si può basare sulla ricetta *B* = "trovare la porta di questa stanza", dato che, una volta trovata la porta, è relativamente semplice uscire. Poiché sappiamo come risolvere *A* facilmente se sappiamo risolvere *B*, questo mostra che il problema *A* è *meno difficile* che *B*. È in questo modo che gli informatici dimostrano che un problema *A* non è più difficile di un altro *B*: supponendo che si conosca una ricetta per *B* e scrivendo una ricetta efficiente per *A* che usa quella di *B*. In questo caso si dice che la soluzione di *A si riduce* a quella di *B*.

Ora il nostro problema di disporre gli ospiti (*Max-Cut*) è parte di quel gruppo di problemi ai quali è possibile ridurre qualsiasi altro problema problema "*ragionevole*": se mi fornirete una ricetta efficiente per il nostro problema Max-Cut, riuscirò a scrivere una ricetta per risolvere qualsiasi altro problema "ragionevole". Ciò significa che tra i problemi che possono essere risolti in un tempo "ragionevole", Max-Cut è uno dei più difficili! Si dice che Max-Cut è *NP-completo* (vedi il riquadro *I problemi più difficili: i problemi NP-completi*).

P =? NP: *la meccanizzazione dell'intuizione è possibile?*

Una questione fondamentale sollevata dagli informatici è espressa da

$$P =? NP$$

che chiede se c'è un metodo efficiente per risolvere tutti i problemi per i quali si può efficacemente *verificare* una soluzione una volta che la si conosce.

P è l'insieme di problemi per i quali esiste una costruzione efficiente della soluzione (vedi *Metodi efficienti e inefficienti* nel riquadro precedente). Per molte questioni importanti (in particolare per l'industria), nessuno conosce, attualmente, algoritmi efficienti. Ad esempio, sui computer di oggi, nessun metodo conosciuto riesce a trovare in modo efficiente il percorso più breve passante per tutte le città di un paese che termina al punto di partenza con più di 1.500 città.

Tuttavia, per la maggior parte di questi problemi importanti, è molto facile verificare efficientemente che una soluzione proposta da qualcuno sia effettivamente corretta o meno (questi problemi sono chiamati NP). Ad esempio, non è noto alcun metodo efficiente per trovare una colorazione di una carta con tre colori in modo che due nazioni vicine abbiano colori diversi, ma questa proprietà è molto facile da verificare per una data colorazione. Ciò significa che, per i problemi NP, se si ha una buona *intuizione* di quella che dovrebbe essere la soluzione, si può efficientemente verificare che la soluzione sia effettivamente corretta.

La questione P=?NP chiede dunque se tutti i problemi NP ammettono una soluzione efficiente. Essa riassume così la questione filosofica "l'intuizione è meccanizzabile"? Questa domanda figura nella lista dei sette problemi matematici per il prossimo millennio, selezionati dall'Istituto Clay nel 2000, per i quali la soluzione sarà ricompensata da un milione di dollari. L'impatto della risposta, ancora oggi sconosciuta, andrebbe assai oltre il contesto informatico.

Il grande problema aperto in informatica è: sarà possibile un giorno trovare una soluzione efficiente per il gruppo dei problemi più difficili? La maggior parte dei ricercatori pensa di no, ma non abbiamo ancora una risposta chiara per il momento (vedi il riquadro *P =? NP: la meccanizzazione dell'intuizione è possibile?*).

I problemi più difficili: i problemi NP-completi

La maggior parte dei problemi computazionali da risolvere nella vita reale appartengono alla classe NP, quelli in cui l'intuizione può essere d'aiuto (vedi riquadro *P =? NP: la meccanizzazione dell'intuizione è possibile?*).

Taluni problemi NP sono più difficili o più importanti di altri, nel senso che sono quelli in cui si concentrano tutte le difficoltà dei problemi NP? Nel 1971, da una parte e dall'altra del muro di Berlino, Stephen A. Cook e Leonid A. Levine hanno dimostrato che c'era effettivamente un problema NP che era, in realtà, più difficile di tutti gli altri, nel senso che tutti gli altri problemi NP si potevano riportare a quello: se vi fosse un metodo efficiente per risolvere quello, saremmo in grado di trasformarlo in un metodo di risoluzione efficiente per tutti i problemi NP!

Tale problema è chiamato NP-*completo*. Richard M. Karp dimostrò l'anno successivo che, in effetti, moltissimi problemi NP, che a priori non hanno nulla a che fare l'uno con l'altro (colorare una mappa, il commesso viaggiatore ...), sono tutte NP-completi! La risoluzione di uno solo di questi condurrebbe alla risoluzione di tutti gli altri! Questo è, per esempio, il caso del problema Max-Cut qui studiato.

La maggior parte dei numerosi problemi studiati in informatica sono in P oppure sono NP-completi, e sono rari quelli che sfuggono a questa dicotomia – tra i quali, però, c'è un caso emblematico: la fattorizzazione di grandi numeri, attualmente pietra angolare della sicurezza di Internet con protocollo RSA (vedi articolo pag. 83).

Nonostante degli sforzi significativi, nessun progresso decisivo che possa indicare l'esistenza o l'impossibilità di una risoluzione efficiente di uno qualsiasi di questi problemi NP-completi è stato ancora compiuto. Data la diversità di questi problemi, molti ricercatori ritengono che non esista un modo efficiente per risolvere i problemi NP-completi.

Frolliamo il nostro problema ...

Fino a quando tale questione non verrà risolta, sarà comunque necessario trovare il modo di disporre i nostri ospiti. Se il problema è troppo difficile, può darsi che rilassando alcuni vincoli, si possa renderlo più abbordabile! Ad esempio, si potrebbe essere soddisfatti di una ricetta che garantisca di tagliare almeno il 99% del numero di bordi tagliati dalla soluzione ottimale. Purtroppo, si può dimostrare che avvicinarsi a 16/17 ovvero 94,1% del numero dei bordi del taglio ottimale è anch'esso parte del "club

dei problemi più difficili” e quindi, probabilmente, non ammette una ricetta efficiente.

Se il problema è troppo difficile, può darsi che rilassando alcuni vincoli, si possa rendere più abbordabile!

Cerchiamo di indebolire ulteriormente i vincoli. Per questo, usiamo una forma originale di programmazione: la programmazione geometrica. Uno dei problemi principali con l’approccio fisico menzionato all’inizio è che le sfere e le molle hanno difficoltà a muoversi in un corridoio troppo stretto e per interferire tra loro, il che complica notevolmente la convergenza verso uno stato di equilibrio ottimale. Se diamo più spazio, è possibile che si possa garantire una convergenza efficiente del dispositivo ad un equilibrio ottimale e persino certificare la qualità della soluzione prodotta.

Una grossa cipolla e qualche chiodo di garofano

Ecco come procederemo. Anziché con delle biglie, rappresenteremo gli n invitati con altrettanti vettori di lunghezza 1 (detti anche *unitari*) applicati ad una origine comune, in uno spazio n-dimensionale (significa solo che abbiamo scelto di dare n componenti a ciascun vettore, affinché essi abbiamo abbastanza spazio per muoversi – la nostra biglia di prima aveva una sola dimensione per muoversi: la sua posizione nel corridoio, questo era troppo restrittivo). Come per le biglie, vogliamo distanziare due vettori corrispondenti agli ospiti collegati da un arco di inimicizia. Invece di collegare questi vettori da molle compresse, cercheremo di

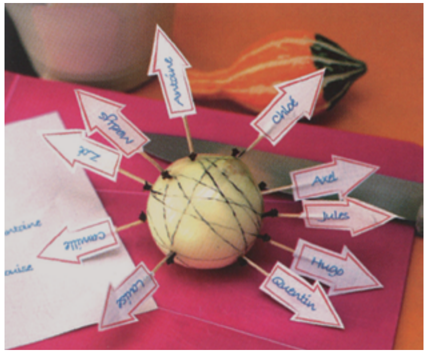


Figura 2. Una proiezione tridimensionale su di una cipolla dei nostri n vettori n-dimensionali nella configurazione ottimale determinata dai nostri vincoli sui prodotti scalari.

puntarli in direzioni opposte. Per questo, useremo il fatto che il *prodotto scalare* di due vettori è positivo se puntano nella stessa direzione e negativo nel caso contrario, anzi più negativo quando puntano in direzioni opposte. Cerchiamo allora di disporre i vettori in modo da *minimizzare la somma dei prodotti scalari* tra i vettori corrispondenti agli ospiti legati da un arco d’ostilità. Ora, si dà il caso che si abbia a disposizione una ricetta (un programma) per risolvere *efficientemente* tali vincoli sui prodotti scalari tra n vettori n-dimensionali: è la “programmazione semi-definita” che generalizza la “programmazione lineare”.

Tagliamo nel mucchio

Grazie alla nostra ricetta, ora abbiamo ottenuto una configurazione dei nostri n vettori unitari n-dimensionali che minimizza la somma dei prodotti scalari dei vettori corrispondenti agli ospiti nemici. Per dividerli in due gruppi, a questo punto è sufficiente tagliare con un piano (in realtà, un *iperpiano* in dimensione n) per l'origine dei vettori in una direzione scelta *a caso*: gli ospiti saranno nell'uno o nell'altro tavolo a seconda che i loro vettori cadano nell'uno o nell'altro lato del piano. La scelta di una direzione casuale per tagliare è un ingrediente essenziale della ricetta. È anche possibile procedere anche senza una scelta casuale, ma è più complicato e paradossalmente si basa in modo essenziale sull'analisi del metodo casuale. Michel Goemans e David Williamson hanno dimostrato nel 1994 che questa soluzione interseca sempre almeno l'87,8% del numero ottimale di archi. In questo modo ho raggiunto il mio obiettivo con un margine di errore relativamente piccolo, garantito essere inferiore al 12,2%.

Tagliare sempre

Esistono molte varianti di questo problema: possiamo provare a tagliare in due parti uguali, o in k pezzi, o aggiungere pesi sugli archi per quantificare le inimicizie tra gli individui: la soluzione proposta precedentemente può essere adattata di conseguenza.

I problemi di taglio nei grafi hanno molte applicazioni nella pratica. Si può citare, ad esempio, il rilevamento delle *comunità* nei grafi, le cui applicazioni sono molteplici: in sociologia, dove possiamo isolare gruppi di amici (con poca inimicizia tra di loro); nel marketing, in cui è possibile suggerire ad una persona un acquisto tramite la sua appartenenza a un profilo di consumo; nella progettazione di microprocessori, in cui i componenti possono essere raggruppati per minimizzare i cablaggi che si incrociano …

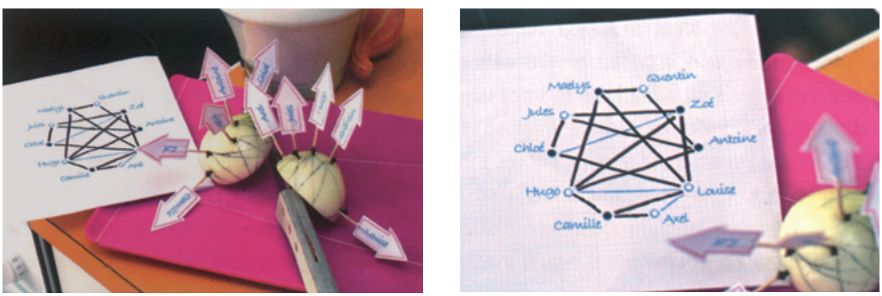


Figura 3. Una volta tagliata seguendo una direzione casuale, otteniamo la partizione degli ospiti qui sopra, che taglia in modo efficiente molte inimicizie (in nero), che si può certificare essere in media superiore all'87,8% di quella ottimale!

In conclusione, è possibile vedere un bel legame tra una geometria di natura *continua* (posso spostare i miei vettori a poco a poco nello spazio) e un problema puramente discreto (le soluzioni possibili sono distintamente separate tra loro, non ci sono intermezzi possibili tra piazzare un invitato su un tavolo o su un altro). Anche se la tecnologia dell'informazione è il regno del discreto (tutto è codificato come numeri interi rappresentati in memoria da sequenze finite di bit, 0 o 1), sembra che le interazioni tra continuo e discreto, non si limitino all'ottenimento di buone approssimazioni. In effetti, alcuni risultati recenti stabiliscono connessioni interessanti tra le congetture circa la difficoltà di risolvere informaticamente certi problemi e certe congetture puramente geometriche.

Questi problemi hanno molte applicazioni nella vita pratica: nel marketing, in cui è possibile suggerire ad una persona un acquisto tramite la sua appartenenza a un profilo di consumo; nella progettazione di microprocessori, in cui i componenti possono essere raggruppati per minimizzare i cablaggi che si incrociano ...

Ma mi rendo conto che i miei ospiti stanno arrivando ed è il momento di sedersi a mangiare, prima di rischiare una indigestione matematico-informatica!

Mantenere il controllo ... ... con l'ausilio della matematica

Karine Beauchard, *ricercatrice CNRS presso l'École polytechnique*
Jean-Michel Coron, *professore presso Université Pierre et Marie Curie*
Pierre Rouchon, *professore presso Mines-ParisTech*

Vincolare la traiettoria di un satellite, regolare la temperatura della propria abitazione, stabilizzare il livello dell'acqua di un canale ... le situazioni che richiedono che debba essere controllato un dato, una quantità, una posizione, sono onnipresenti. Questi temi sono oggetto di una ricca teoria matematica: la teoria del controllo.

Posiamo una scopa verticalmente, capovolta, su una superficie piana (la punta della scopa è arrotondata). La posizione in cui la scopa è verticale, immobile, è una posizione di equilibrio: in teoria, quando la scopa è in questa configurazione, vi rimane. In pratica, questo non è ciò che constatiamo (la scopa si allontana dalla verticale e finisce per cadere), perché la scopa non è esattamente sulla posizione di equilibrio (non è né perfettamente verticale, né perfettamente immobile). Così, piccoli disturbi producono grandi effetti: anche se all'inizio dell'esperimento la scopa è vicina alla posizione di equilibrio, finisce per allontanarsene notevolmente. Si dice allora che l'equilibrio è instabile.

Adesso, posiamo la scopa sul nostro indice. Spostando il dito abilmente, si può impedire che la scopa cada, anche se questo richiede un certo allenamento. Nel gergo della teoria del controllo, il movimento del nostro dito è un *comando* che agisce sul *sistema* (la scopa). Infatti, il nostro dito esercita una retroazione sulla scopa, chiamata anche *feedback*, in modo da rendere stabile un equilibrio instabile. Questo feedback si basa sul fatto che i nostri occhi vedono la scopa cadere e quindi, se la scopa inizia a pendere a destra, muoveremo il nostro indice a destra, e questo tanto più rapidamente quanto più velocemente pende la scopa. Così la posizione del nostro indice, il comando, dipende ad ogni istante da una combinazione di posizione e velocità della scopa: questo è un feedback che stabilizza la scopa intorno al suo equilibrio verticale instabile.

Il nostro dito esercita sulla scopa una retroazione, nota anche come "feedback", in modo da rendere stabile un equilibrio instabile.

Questo problema di stabilizzazione della scopa è un tipico problema detto di *controllo*. La teoria del controllo studia i sistemi su cui si può agire tramite un comando. Il comando è qui utilizzato per portare il sistema da un dato stato iniziale ad uno stato finale desiderato nonostante la presenza di disturbi. Il problema della stabilizzazione è una questione centrale della teoria del controllo.

Gli orologi ad acqua di Alessandria

Storicamente, uno dei feedback più antichi mai costruiti dall'uomo è quello inventato da Ctesibio, un ingegnere greco di Alessandria del III secolo a.C, per orologi ad acqua o clessidre, il cui scopo è quello di misurare il tempo che passa, un problema che risale ai primordi della civiltà. Il principio della clessidra è il seguente: un primo contenitore, forato sul fondo, è riempito con acqua. L'acqua fluisce attraverso il foro in un altro contenitore. Si può quindi misurare il tempo misurando l'altezza dell'acqua nel secondo contenitore. Questo metodo è stato utilizzato dagli egizi dalla metà del secondo millennio a.C. Il problema di questo dispositivo è che è difficile garantire un flusso costante del primo contenitore al secondo: più acqua c'è nel primo contenitore, maggiore è la velocità, come mostrato in figura 1. Quindi, il flusso dell'acqua diminuisce nel tempo e il livello dell'acqua nel secondo contenitore non è proporzionale al tempo trascorso.

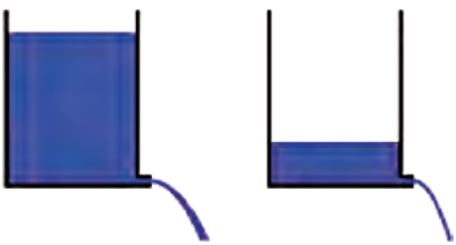


Figura 1. Quando il livello scende, la portata diminuisce.

Per superare questo problema, Ctesibio ha introdotto nel III secolo a.C. il dispositivo mostrato in figura 2. Adesso ci sono tre contenitori; l'acqua fluisce dal primo contenitore al secondo e poi dal secondo fino al terzo. È nel secondo contenitore che si trova il feedback. Quando il livello dell'acqua nel contenitore 2 è troppo alto, il galleggiante rosso tocca il coperchio verde. Questo impedisce all'acqua di passare dal contenitore 1 al contenitore 2. Viceversa, se il livello dell'acqua nel contenitore 2 è bassa, il galleggiante rosso non ostruisce il flusso d'acqua dal contenitore 1 al contenitore 2. L'apparecchio è costruito in modo che la portata dal contenitore 1 al contenitore 2 sia maggiore di quella dal recipiente 2 al recipiente 3, quindi il livello dell'acqua aumenta nel contenitore 2, fino a che il galleggiante rosso non sia di nuovo troppo alto. In questo modo si mantiene un livello quasi costante d'acqua nel contenitore 2, che assicura un flusso costante dal contenitore 2 al contenitore 3. Il livello dell'acqua nel recipiente 3 dà così una misura relativamente buona del tempo trascorso.

Il primo regolatore industriale

Il primo regolatore utilizzato industrialmente è il regolatore di James Watt (1736-1819), che serve a regolare la velocità di rotazione di un motore a vapore. Lo si può osservare in figura 3 e nello schema 4. Il suo funzionamento è il seguente. Quando la macchina è in funzione trasmette (con un sistema non mostrato) il suo movimento all'asse del regolatore, che ruota attorno alla verticale ad una velocità proporzionale alla velocità di rotazione della macchina. La struttura del regolatore è solidale con il suo asse e ruota quindi alla stessa velocità. Quando la velocità di rotazione aumenta, le due sfere, sotto l'azione della forza centrifuga, si allontanano dall'asse verticale e il manicotto rosso si alza. Questo manicotto rosso attraverso un insieme di aste non mostrato, agisce sulla valvola di immissione del vapore nella macchina. Pertanto, se la macchina ruota più velocemente di quanto desiderato, il manicotto si solleva: questo provoca una riduzione della fornitura di vapore, e la macchina inizia a girare più lentamente. Viceversa, se la macchina ruota più lentamente di quanto desiderato, il manicotto si muove verso il basso: questo provoca una maggiore apertura della valvola di immissione, aumentando l'erogazione di vapore, e la macchina inizia a girare più velocemente. Il regolatore di Watt esercita quindi un feedback sul motore a vapore. Questa azione dipende, in ogni momento, dallo stato della macchina a vapore.

È stato solo 80 anni più tardi che la prima analisi matematica di questo tipo di controllo è stata effettuata da James Clerk Maxwell (1831-1879). Il regolatore di Watt viene modellato in questo contesto da un'equazione differenziale.

Nel mondo di oggi, i regolatori sono onnipresenti. Ad esempio, il termostato di una casa stabilizza la temperatura della casa vicino al

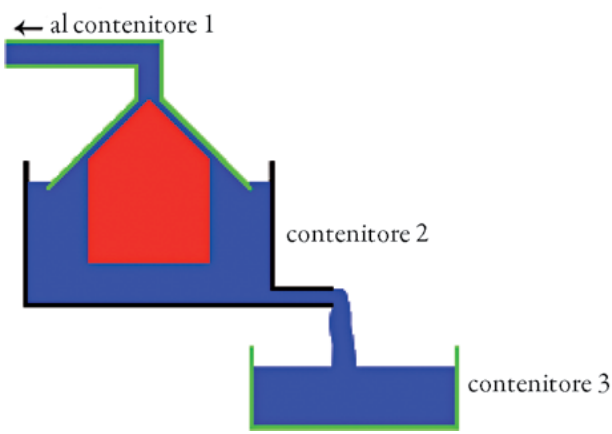


Figura 2. Grazie al galleggiante rosso, il flusso è adesso costante.

Figure 3. Questo regolatore è installato sulla celebre macchina a vapore "Lap" della società Boulton & Watt che risale al 1788 e si trova attualmente presso lo Science Museum a Londra.

valore impostato dall'utente: il termostato regola in tempo reale la posizione della valvola, che determina se l'acqua nel circuito dei radiatori passa dalla caldaia (valvola aperta) o circola nel circuito senza essere riscaldata (valvola chiusa). Le sospensioni attive delle auto effettuano una stabilizzazione dell'abitacolo attraverso un feedback. I droni volanti sono dotati di feedback che li stabilizzano e li rendono più facili da guidare, ecc.

Anche se James Watt è riuscito a regolare la sua macchina a vapore senza analisi matematica, i due esempi recenti di regolatori che ora vedremo si basano in modo cruciale sulla matematica: il primo stabilizza i corsi d'acqua navigabili, il secondo gli stati quantici fragili.

Le equazioni di Saint-Venant per gli invasi

I canali navigabili sono costituiti da gore, vale a dire, da una serie di piccoli canali separati da paratoie mobili che permettono di agire in tempo reale sul livello dell'acqua, e da invasi per il passaggio delle barche (vedi figura 5). In questi canali navigabili è importante regolare sia il livello dell'acqua sia la velocità del flusso. Perché? In primo luogo, dobbiamo permettere ai portacontainer di calcolare il carico massimo che possono montare a bordo senza correre il rischio di toccare il fondo, il che rende necessario regolare il livello dell'acqua con precisione, dato che i canali sono spesso poco pro-

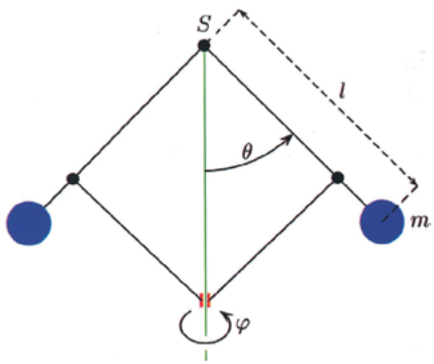


Figura 4. Schema del regolatore di Watt.

fondi. Inoltre è necessario garantire continuità all'approvvigionamento alle centrali elettriche e alle industrie che consumano acqua situate sulle rive, il che richiede una regolazione del flusso.

Il controllo delle paratoie di una gora avviene automaticamente. Il processo è gestito da un algoritmo basato su un modello di controllo che descrive il comportamento dell'acqua negli invasi.

Il controllo delle paratoie oggigiorno avviene automaticamente. Il processo è gestito da un algoritmo basato su un modello di controllo che descrive il comportamento dell'acqua negli invasi. Il modello più diffuso è quello delle equazioni di Saint Venant. La sua origine risale al XIX secolo. Adhemar-Jean-Claude Barré de San Venant è stato uno studente della École Polytechnique e, successivamente, ha insegnato matematica presso la École nationale des ponts et chaussées. Pubblicò il suo modello nel 1871, frutto di uno studio sui fiumi in piena e sul movimento delle maree negli estuari. Si tratta di equazioni alle derivate parziali che traducono la conservazione della massa e della quantità di moto in un fluido.

Per progettare leggi di controllo che regolino il comportamento delle paratoie, si utilizza una quantità rappresentativa del comportamento generale dell'invaso: una *funzione di Lyapunov*, cioè una funzione che raggiunge il suo minimo quando il sistema è nello stato di equilibrio desiderato. Questo metodo di stabilizzazione mediante funzioni di Lyapunov è molto comune in teoria del controllo. Ha il vantaggio

Figura 5. Paratoie mobili per i canali di controllo. A destra, la paratoia è fuori dall'acqua per una operazione di manutenzione.

di condurre a dei comportamenti affidabili nei confronti delle incertezze dovute alla modellizzazione e ai disturbi esterni. Inoltre, queste leggi di controllo sono, nel nostro esempio, facili da implementare in pratica.

Questo metodo di analisi e progettazione del controllo dei corsi d'acqua è stato utilizzato, in particolare, per l'impostazione delle paratoie del Sambre da parte dell'amministrazione delle vie idrauliche del Ministero delle Infrastrutture e dei Trasporti della Regione Vallona in Belgio. Si sono potuti constatare dei guadagni di prestazioni notevoli: l'ampiezza delle variazione del livello è stata dimezzata e le perturbazioni del flusso vengono ammortizzati due volte più velocemente.

Stabilizzare gli stati quantici di luce usando la probabilità

L'esperienza che segue è la prima realizzazione sperimentale di un feedback per un sistema quantistico. Si tratta di un passo significativo verso il computer quantistico, per cui la protezione di stati quantici attraverso un feedback sarebbe un metodo efficace per combattere i disturbi legati all'ambiente (decoerenza). L'obiettivo di questo esperimento è di stabilizzare certi stati quantici della luce. Questi stati comportano un numero intero e ben definito di fotoni. Questi sono fragili e difficili da osservare perché molto diversi da quelli della luce che ci circonda. L'obiettivo è di mantenere costante tale numero di fotoni. Come mostrato in modo schematico in figura 6, dobbiamo superare una difficoltà aggiuntiva, in quanto su scala quantistica, la misurazione modifica il sistema. In questo esperimento, i fotoni sono confinati fra due specchi superconduttori uno di fronte all'altro, formando così una cavità aperta ai lati, come mostrato in figura 7. La misurazione si effettua mediante degli atomi che passano uno dopo l'altro attraverso la cavità, interagendo con i fotoni, e misurati in uscita. Per il comando, si dispone di una sorgente di luce adatta, usata tra il passaggio di un atomo e l'altro. (Una descrizione molto più dettagliata di questo esperimento si trova in particolare sul sito http://www.lkb.ens.fr/Un-asservissement-quantique/ dove vengono anche presentate delle animazioni basate sui dati sperimentali.)

I principi della meccanica quantistica conducono ad una descrizione probabilistica molto precisa del modello che collega i segnali di comando e quelli di misura. Questa descrizione matematica, anche se relativamente complessa, permette di determinare un feedback stabilizzante che ancora una volta si basa su una funzione di Lyapunov.

> *Anche se i modelli utilizzati per i canali di navigazione e gli stati quantici di luce sono di natura molto diversa, i loro controlli si basano su metodi matematici analoghi, che risalgono al lavoro fondamentale di Alexander Lyapunov sulla stabilità*

Così, anche se i modelli utilizzati per i canali di navigazione e quelli degli stati quantici di luce sono di natura molto diversa, essendo il primo determinista e con *tempo continuo* e il secondo probabilistico e con *tempo discreto*, i loro controlli si basano sui metodi matematici simili la cui origine risale ai lavori fondamentali sulla stabilità di Alexander Lyapunov.

Nei problemi di controllo, il progresso non viene quindi solo puramente da invenzioni tecniche o sperimentali. Nasce anche dalle ricerche astratte della teoria matematica del controllo. Gli strumenti matematici ivi utilizzati sono molteplici (equazioni differenziali, equazioni differenziali alle derivate parziali, processi stocastici, ...) e la ricerca spesso si svolge nell'interfaccia con altre discipline (meccanica dei fluidi, idraulica, meccanica quantistica, ottica quantistica, ...).

Figura 6: Gli stati quantici da stabilizzare sono degli stati con un numero intero ben definito di fotoni. Questi rimbalzano tra i due specchi. La difficoltà per il feedback, qui rappresentato metaforicamente dalle mani del giocoliere, è che l'osservazione di questi fotoni da parte del giocoliere inevitabilmente li perturba. Questa retroazione dovuta alla misurazione deve essere messa in conto nella progettazione del feedback stabilizzante.

Ringraziamenti

Gli autori ringraziano Michel Brune per le figure 6 e 7 che illustrano il controllo degli stati quantici.

Bibliografia

Coron J.-M., d'Andréa-Novel B., Bastin G., (marzo 2008). *Penser globalement, agir localement*. La Recherche, n° 417, p. 82-83.

Mayr O., (1970). *The origins of feedback control*. The M.I.T. Press (Cambridge Massachusetts, and London, England), p. vii +151.

Sayrin C., Dotsenko I., Zhou X., Peaudecerf B., Rybarczyk T., Gleyzes S., Rouchon P., Mirrahimi M., Amini H., Brune M., Raimond J.-M., Haroche S., (settembre 2011). *Real time quantum feedback prepares and stabilizes photon number states*. Nature, vol. 477, p. 73-77.

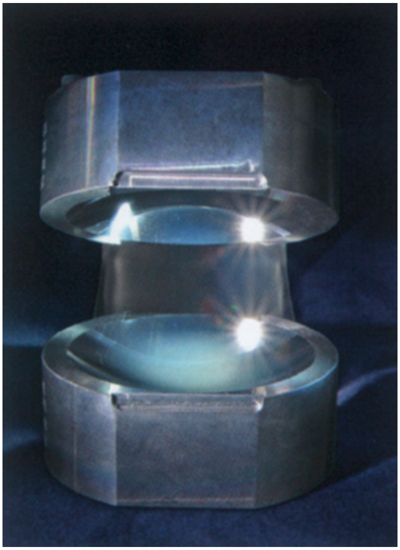


Figura 7. La luce da controllare è un campo elettromagnetico di una frequenza vicina ai 10 GHz, confinata tra due specchi superconduttori contrapposti come nella figura. In questo modo questi formano una cavità aperta ai lati. La misura si basa su degli atomi (non mostrati in figura) che attraversano orizzontalmente la cavità a metà strada tra i due specchi. Questi atomi interagiscono con il campo elettromagnetico e vengono misurati all'uscita. Per il comando, si dispone di una sorgente elettromagnetica vettoriale regolabile (non mostrata in figura) che, dopo ogni atomo, illumina la cavità come un flash.

Il teorema di Green-Tao e altri segreti dei numeri primi

Michel Waldschmidt, *professore emerito presso l'Università Pierre et Marie Curie*

Anche se i matematici se ne sono interessati fin dai tempi antichi, i numeri primi continuano ad affascinare. Sommandoli o sottraendoli tra loro, si trova una miniera di problemi, alcuni dei quali sono rimasti aperti a lungo o restano ancora irrisolti.

Ottenere tutti i numeri interi positivi utilizzando soltanto i simboli + e 1 è facile: basta scrivere qualsiasi intero $n > 0$ nella forma $1 + 1 + \ldots + 1$ con n volte il simbolo 1 e (n-1) volte il simbolo + .

Sostituiamo l'addizione con la moltiplicazione e il simbolo + con il simbolo × (o •) . Per scrivere tutti i numeri interi positivi come prodotto, abbiamo quindi bisogno di *mattoni elementari* che svolgono il ruolo che 1 ha per l'addizione. Questi mattoni sono i *numeri primi*, quelli che non possono essere scritti come il prodotto di numeri interi più piccoli. Ogni numero intero maggiore o uguale a 2 è il prodotto, in modo univoco (a meno di permutazioni) di numeri primi: questo è quello che si chiama il *teorema fondamentale dell'aritmetica*. Per esempio $4200 = 2 \times 2 \times 2 \times 3 \times 5 \times 5 \times 7$.

Numeri grandi a piacere

La lista dei numeri primi inizia con 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, ...
L'inizio di questa lista si trova nell'enciclopedia delle sequenze dei numeri interi, che è una miniera di informazioni per questo tipo di questioni. I numeri superiori a 1 che non sono numeri primi sono chiamati *composti*. La loro lista inizia con 4, 6, 8, 9, 10, 12, 14, 15, 16, ...

Questo secondo elenco contiene tra l'altro tutti i numeri pari a partire da 4. È facile scrivere un numero composto grande quanto vogliamo. Invece, la scrittura (ad esempio nella forma decimale) di un numero primo grande a piacere è una operazione che non sappiamo effettuare.

Il più grande numero primo noto in modo esplicito (dal 2 Maggio 2013) è $2^{57\,885\,161}$-1 che ha 17 425 170 cifre in base dieci. Se scrivessimo tutte queste cifre con due cifre per centimetro, la scrittura del numero si estenderebbe per oltre 87 km!

Eppure sappiamo da Euclide, che la lista dei numeri primi non si ferma mai: *esiste un numero infinito di numeri primi*. I numeri primi sono quindi un esempio di successione che sappiamo essere infinita, ma per la quale non sappiamo esplicitare degli elementi grandi a piacere.

Proprietà additive sorprendenti

I numeri primi sono stati introdotti in aritmetica per la moltiplicazione, come abbiamo appena visto. Studiare le loro proprietà additive può sembrare una strana idea. Tuttavia, questo ci conduce ad alcuni problemi profondi, il cui studio ha portato a teorie molto elaborate e di grande ricchezza, che richiedono l'utilizzo di sofisticati strumenti matematici. Ma la cosa più affascinante di questa teoria è che essa continua ad essere sviluppata. Essa ha permesso di risolvere un grande numero di questioni che sono rimaste aperte per lungo tempo, ma non è ancora abbastanza potente per rispondere ad alcune domande che tuttavia sono facili da porre.

Sommiamo due numeri primi tra loro; si ottengono i seguenti valori:

p	2,	3,	5,	7,	11,	13,	17,	19,	23,	29,	31,	37...
$p+2$	4,	5,	7,	9,	13,	15,	19,	21,	25,	31,	33,	39...
$p+3$	5,	6,	8,	10,	14,	16,	20,	22,	26,	32,	34,	40...
$p+5$	7,	8,	10,	12,	16,	18,	22,	24,	28,	34,	36,	42...
$p+7$	9,	10,	12,	14,	18,	20,	24,	26,	30,	36,	38,	44...
$p+11$	13,	14,	16,	18,	22,	24,	28,	30,	34,	40,	42,	48...
...	...											

Gli unici numeri dispari che possiamo trovare in questa tabella sono riportati nella prima riga (quella che dà i $p+2$) e sulla prima colonna (quella che dà i $2+p$, poiché la tabella è simmetrica). Infatti, la somma di due numeri dispari è necessariamente un numero pari. Dato che 2 è l'unico numero primo pari, le uniche somme di due numeri primi dispari sono quelli che coinvolgono l'aggiunta di 2 ad un primo dispari; tutte le altre somme sono numeri pari.

Una cosa notevole è che sembra si possano ricavare da questa tabella tutti i numeri pari maggiori di 4. Molti numeri pari si possono esprimere in vari modi, ad esempio,

$10 = 5+5 = 3+7;\ 14 = 7+7 = 3+11...$

Che tutti i numeri maggiori o uguali a 4 possano essere scritti come somma di due numeri primi è un problema ancora aperto.

Che tutti i numeri pari maggiori o uguali a 4 possano essere scritti come somma di due numeri primi è chiamato problema binario di Goldbach; è stato proposto da Eulero in risposta ad una lettera di Goldbach risalente al 1742, in cui Goldbach suggeriva che qualsiasi numero intero maggiore o uguale a 7 è la somma di tre primi. Quest'ultimo risultato, più debole del problema binario di Goldbach, è stato annunciato da Harald Helfgott nel maggio 2013. Ma il problema binario di Goldbach è ancora aperto ed è stata la fonte di molte ricerche per quasi tre secoli. Questo è un problema difficile, che ha dato luogo a teorie sofisticate che hanno altre applicazioni. Possiamo citare ad esempio il problema della teoria additiva dei numeri, che studia gli interi appartenenti ad un insieme dato (qui i numeri primi).

La differenza tra due numeri primi: una fonte di domande appassionanti

Un modo diverso per accoppiare i numeri primi rispetto all'addizione è quello di studiare la differenza tra numeri primi consecutivi.

La lista dei numeri primi è p_1, p_2, p_3, ...) = (2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37,...) e la lista ($p_n - p_{n-1}$) delle differenze tra due numeri primi consecutivi inizia con 1, 2, 2, 4, 2, 4, 2, 4, 6, 2, 6 ...

È facile vedere che a parte il primo valore, 1, tutti gli altri sono numeri pari: questo ancora una volta è dovuto al fatto che 2 è l'unico numero primo pari. Ci sono un infinità di 2 in questa lista? Non lo sappiamo ancora; è la congettura dei primi gemelli, che afferma: *c'è un'infinità di numeri primi p con la proprietà che $p+2$ è anch'esso primo*.

Sappiamo che le differenze tra due numeri primi consecutivi della successione di numeri primi non è limitata: ci sono "buchi" grandi a piacere (lo studio dei "buchi" nella successione dei numeri primi risale a Legendre e Gauss). Sappiamo anche stimare l'ordine di grandezza media di tali buchi, grazie al Teorema dei numeri primi, dimostrato nel 1896 da Hadamard e da la Vallée Poussin. La leggenda racconta che il matematico che riuscisse a dimostrare questa congettura diventerebbe immortale. In qualche modo Jacques Hadamard (1865-1963) e Charles de la Vallée Poussin (1866-1962) hanno lasciato il loro nome nella storia con il loro teorema, ma hanno anche quasi avverato la previsione, dal momento che sono morti dopo quasi un secolo di vita.

Di recente, sono stati compiuti nuovi progressi sulla questione dei buchi tra due numeri primi. Il postulato di Bertrand, dimostrato da Chebyshev nel 1850, afferma che tra un numero intero e il suo doppio c'è sempre un numero primo.

Anziché raddoppiare, cioè usare il fattore 2, possiamo prendere qualsiasi fattore $c > 1$: per n sufficientemente grande, vi è sempre un primo p tra n e cn. Questo dà già informazioni sulle differenze tra elementi consecutivi della successione dei numeri primi. Il teorema dei numeri primi implica più precisamente che la differenza tra l'n-esimo numero primo p_n e il successivo p_{n+1} è nell'ordine di grandezza di $\log p_n$, il logaritmo naturale di p_n.

Nel 2013, Yitang Xang ha dimostrato che esistono infinite coppie di numeri primi la cui differenza è inferiore a 70 milioni.

Ciò è ancora lontano da quello che ci aspetteremmo pensando alla congettura dei primi gemelli esposta sopra, secondo la quale la differenza tra due numeri consecutivi nella lista dei numeri primi dovrebbe essere 2 infinite volte. Ma questo è un considerevole miglioramento rispetto a quello che sapevamo prima.

Progressioni aritmetiche

Lo studio delle differenze tra numeri primi può essere generalizzato in un altro modo, studiando progressioni aritmetiche composte da numeri primi. Una successione di numeri forma una *progressione aritmetica* se la differenza tra due termini consecutivi della successione è costante. Questo valore è la *ragione* della progressione aritmetica.

Per esempio la successione 5, 11, 17, 23, 29 è una progressione aritmetica di 5 termini di ragione 6.

La successione 7, 37, 67, 97, 127, 157 è una progressione aritmetica di 6 termini di ragione 30 e la successione 199, 409, 619, 829, 1039, 1249, 1459, 1669, 1879 , 2089 è una progressione aritmetica di 10 termini di ragione 210. In questi esempi, i termini di queste successioni sono tutti dei numeri primi.

Esistono progressioni aritmetiche lunghe a piacere costituite unicamente da numeri primi?

La progressione aritmetica più lunga nota esplicitamente composta solo da numeri primi contiene 26 termini, la cui ragione è il prodotto di 23 681 770 per 223 092 870; è la successione 43 142 746 595 714 191 + 223 092 870 $\cdot$ 23 681 770 $\cdot\, n$ con n che varia da 0 a 25. Essa è stata scoperta da Bonoît Perichon.

Esistono progressioni aritmetiche lunghe a piacere composte unicamente da numeri primi? Per molto tempo questo problema è rimasto aperto, era dunque una sfida in più per i matematici. È stato un lavoro di collaborazione trai matematici Ben Green e Terence Tao che alla fine ha portato alla soluzione nel 2004. Il loro teorema afferma che sì, ci sono progressioni aritmetiche di lunghezza grande a piacere, composte da numeri primi soltanto. La loro dimostrazione non è costruttiva, vale a dire, non ci dice come costruire una tale successione!

Ben Green

Terence Tao

Per approfondire

Tenenbaum G., Mendès-France M., (1997). *Les Nombres premiers*; Presses Universitaires de France (PUF), coll. Que Sais-je?.

Ribenboim P., (2000). *Nombres premiers: mystères et records*; Presses Universitaires de France (PUF), coll. Mathématiques.

Hardy G.H., Wright E.M., (2008). *An Introduction to the Theory of Numbers*; Oxford Science Publications, sixth ed. Oxford: Oxford University Press. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

Su Internet

http://www.mersenne.org/
http://primes.utm.edu/largest.html
https://oeis.org/

Il blog di Terence Tao:
http://terrytao.wordpress.com/

La superconduttività

Sylvia Serfaty, *professore presso Université Pierre et Marie Curie*

La superconduttività, la capacità di un metallo di far passare la corrente elettrica senza perdita di energia, può avere applicazioni sorprendenti. Lo studio di questo fenomeno coinvolge vari settori della matematica, come il calcolo delle variazioni, le equazioni alle derivate parziali, l'analisi asintotica. Ad esso sono associati diversi problemi aperti.

La superconduttività ha appena festeggiato 100 anni. Più precisamente, è stato nel 1911 che il fisico olandese Heike Kamerlingh Onnes ha scoperto che, una volta raffreddato ad una temperatura molto vicina allo zero assoluto, il mercurio perde completamente la sua resistenza elettrica: in altre parole, si osserva circolare su di esso delle correnti elettriche indefinitamente, senza dissipazione d'energia o di perdita di calore.

Levitazione e vorticosità

La superconduttività, che riguarda molti metalli e leghe, provoca altri fenomeni inconsueti, in particolare in risposta ad un campo magnetico. Quando un superconduttore, raffreddato al di sotto della sua *temperatura critica*, è sottoposto ad un campo magnetico esterno, provoca un chiaro effetto respingente. Questo è chiamato *effetto Meissner*.

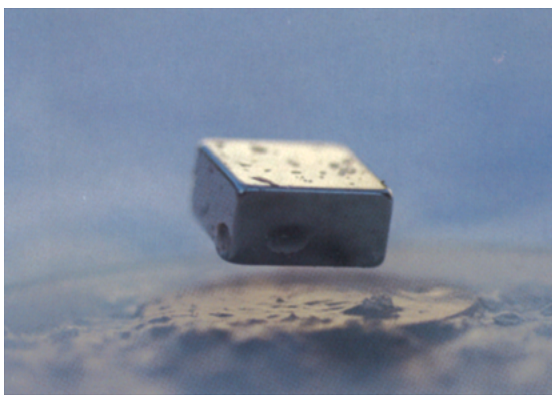


Figura 1. Un superconduttore levita sopra un magnete: l'effetto Meissner. Copyright: J. Bobrov

Una delle conseguenze più sorprendenti di questo effetto è che un magnete posto sopra un superconduttore è sottoposto ad una forza magnetica che lo respinge: si osserva il magnete levitare sopra il superconduttore!

Tuttavia, quando certi materiali superconduttori sono sottoposti ad un campo magnetico più intenso, la situazione cambia: oltre una certa soglia, chiamata *primo campo critico*, il campo magnetico comincia a penetrare in certi punti, e là dove si insinua, si formano dei *vortici*, che sono come dei mini-cicloni, al cui centro si ha lo stato normale (non-superconduttore) circondati dalla fase superconduttrice e attorno ai quali ruotano degli anelli di *corrente superconduttrice.* Più il campo applicato è elevato, più i vortici si infittiscono. All'aumentare del loro numero, si organizzano in perfetti reticoli esagonali a nido d'ape (vedi foto). Tuttavia, superata una certa intensità, chiamata *secondo campo critico*, i vortici divengono così fitti tra loro che non vi è più spazio per la fase superconduttiva, e il materiale torna ad assumere il suo stato normale.

Figura 2. Vortici organizzati in rete di Abrikosov, foto di H.F. Hess et al., Bell Labs, Phys. Rev. Lett. 62, 214 (1989)

Un magnete posto sopra un superconduttore è sottoposto ad una forza magnetica che lo respinge: si osserva il magnete levitare sopra il superconduttore!

Un enorme potenziale in termini di applicazioni

Le applicazioni della superconduttività sono potenzialmente enormi. Infatti, questo fenomeno in linea di principio permette di trasportare correnti molto elevate senza perdita di energia. Già oggi, i treni a levitazione magnetica, come il giapponese Maglev, sono basati sulla superconduttività. Questi treni sono dotati di bobine superconduttrici che permettono loro di scivolare senza attrito, levitando al di sopra di rotaie magnetizzate. I superconduttori sono utilizzati anche per generare campi magnetici di forte intensità, facendo fluire correnti elevate, come ad esempio nel nuovo acceleratore di particelle LHC di Ginevra. Vengono inoltre utilizzati per rilevare campi magnetici molto deboli; questo trova impiego, tra l'altro, nei sottomarini, nelle immagini medicali, nel rilevamento di oggetti nascosti e nei nano-circuiti. Al momento, le limitazioni principali sono dovute al fatto che i superconduttori necessitano di essere dotati di sistemi per il raffreddamento a temperature molto basse, il che è costoso e riduce le loro applicazioni. I fisici in collaborazione con i chimici, continuano ad esplorare la possibilità di sintetizzare delle sostanze chimiche che risultino dei superconduttori a temperature più elevate.

Un po' di fisica quantistica

Cosa dovrebbe succedere all'interno di un materiale che diventa superconduttore a basse temperature? Solo la fisica quantistica può spiegarlo. In un metallo normale, gli elettroni si comportano come onde diffuse su ciascun atomo, indipendenti le une dalle altre. Quando il metallo diventa superconduttore a bassa temperatura, gli elettroni si legano a coppie, dette *coppie di Cooper*. Approssimativamente, queste coppie di elettroni si mettono nello stesso stato quantico a formare un'unica onda, proprio come i pesci in un banco. Questo comportamento collettivo degli elettroni permette loro di passare attraverso il materiale senza essere sensibili agli ostacoli e quindi la resistenza elettrica scompare. La *condensazione* in un unico stato quantico è esattamente ciò che avviene anche per gli atomi nei *superfluidi* (come per l'elio liquido a bassa temperatura) o i condensati di Bose-Einstein, previsti da Einstein e recentemente osservati sperimentalmente. Superfluidi e condensati sono due tipi di liquidi che perdono la viscosità a temperature molto basse, da cui il nome di *liquidi superfluidi*. Quando si mette un liquido superfluido in rapida rotazione, si osserva, come nei superconduttori, la formazione di vortici.

Negli anni '30 il fisico London ha fornito il primo modello descrittivo della superconduttività, ma è solo negli anni '50 che è nata la teoria di Ginzburg e Landau, ad oggi quella più universalmente accettata, che descrive la comparsa della superconduttività e il comportamento dei superconduttori in presenza di campo magnetico. Questa teoria descrive un superconduttore mediante due "funzioni" che indicano il suo stato locale: una *funzione d'onda* Ψ e un campo magnetico A, ciascuno dipendente dal punto del campione nel quale ci poniamo.

A queste due funzioni è associato, attraverso una formula esplicita, un numero unico, dipendente dall'intensità del campo applicato, e che fornisce quella che viene chiamata *l'energia di Ginzburg-Landau*. Gli stati stabili del sistema, quelli che si osservano a riposo, sono quelli per cui l'energia Ginzburg-Landau è minima. Nel 1956, la teoria BCS di Bardeen, Cooper e Schrieffer ha completato la teoria di Ginzburg e Landau, spiegando la superconduttività a partire dal livello quantico, mediante la formazione di coppie di Cooper.

Dove interviene la matematica in tutto questo?

L'energia di Ginzburg-Landau, che è una funzione di stato del sistema, costituisce una sorta di funzione matematica. Inoltre, in modo non così sorprendente, questa è quasi la stessa funzione matematica che descrive l'energia di un superfluido o di un condensato di Bose-Einstein in rotazione. Si può dunque fingere di dimenticare il modello fisico da cui questa energia proviene e guardarla con occhio matematico, studiandola come una funzione matematica, con l'aiuto di dimostrazioni, ottenendo in un solo colpo dei risultati validi in ciascuno di questi tre problemi fisici. Ci si domanda allora: qual è il valore minimo di questa funzione? Quali sono gli stati che raggiungono questo minimo? Come variano a seconda del valore del campo applicato? Possiamo predire l'insorgenza di vortici (caratterizzati come i punti in cui la funzione Ψ vale 0) e le loro posizioni?

Questo tipo di domande si pone per tutti i tipi di energia che provengono dalla fisica, così come dall'ingegneria, dall'economia ecc., dove si cerca di minimizzare certi costi (o, nel caso dell'economia, di massimizzare il guadagno o una *funzione utilità*) e dove si vuole capire quali sono gli stati o le configurazioni ottimali. In matematica, questa è chiamata la teoria del *calcolo delle variazioni*, che risale a Bernoulli, Eulero e Lagrange, che ci dice quando siamo in grado di garantire l'esistenza di un *optimum*, e di descriverlo. Di solito, questo viene caratterizzato mediante una *equazione alle derivate parziali*, che altro non è che una relazione esatta tra la funzione di stato (che dipende dalla posizione del campione) e le sue derivate. I matematici hanno sviluppato da più di un secolo una vasta teoria e una classificazione di queste equazioni differenziali alle derivate parziali. Anche se rimangono molte domande difficili, sappiamo dire molto bene quali di esse hanno soluzione, come si comportano qualitativamente nello spazio e nel tempo, ecc.

I matematici sanno dire, per esempio, come i vortici si disporranno sotto l'influenza di una corrente applicata, problema importante poiché i movimenti dei vortici provocano una dissipazione di energia.

Ma torniamo a quella famosa energia di Ginzburg-Landau: come fare per sapere se questa energia, originariamente proposta senza riscontri sperimentali, costituisce un buon modello per descrivere la superconduttività? Più precisamente, si vuole capire se questa predice matematicamente lo stesso comportamento di quello osservato negli esperimenti. Lo stesso problema si presenta per qualsiasi modello fisico: lo studio matematico dapprima aiuta a validare il modello, successivamente permette di predire il comportamento del sistema fisico con maggiore dettaglio e in quelle situazioni che non sono, ad esempio, accessibili sperimentalmente! In questo modo, il fisico Abrikosov, poco dopo l'introduzione del modello di Ginzburg e Landau, ha fatto dei calcoli, piuttosto formali, che lo hanno portato a dire che il modello dovrebbe avere soluzioni con vortici organizzati periodicamente e a predire che questo dovrebbe essere riscontrato sperimentalmente. I fisici hanno in seguito cercato di osservare tali situazioni e hanno effettivamente scoperto la disposizione dei vortici a reticolo esagonale, come si vede nella foto a pagina 38 (Figura 2). Queste disposizioni sono state chiamate *reti di Abrikosov*.

Rispetto alla fisica e al calcolo simbolico, la matematica apporta tutto il rigore e la precisione del metodo ipotetico-deduttivo. Da questo punto di vista, il modello Ginzburg-Landau è già ampiamente convalidato matematicamente. In effetti, siamo in grado di dimostrare per quali valori (detti critici) del campo magnetico applicato appaiono dei vortici nei minimi dell'energia, quanti ce ne sono, e come si dispongono mediamente. Si ritrova e si precisa in questo modo il comportamento osservato o calcolato dai fisici. I matematici sanno dire, per esempio, come i vortici si disporranno sotto l'influenza di una corrente applicata, problema importante, poiché i movimenti dei vortici provocano una dissipazione di energia.

Analisi asintotica

Nell'analisi matematica della superconduttività intervengono quindi il calcolo delle variazioni, la teoria delle equazioni alle derivate parziali, la *teoria spettrale*, ma anche, e soprattutto, un altro ingrediente chiamato *analisi asintotica*. Nei superconduttori, il comportamento del vortice dipende da una costante K, chiamata parametro Ginzburg-Landau, che è specifico per ogni materiale che compare in modo naturalmente nel calcolo dell'energia Ginzburg-Landau. Infatti, è solo nei materiali per i quali questa costante è sufficientemente grande che i vortici fanno la loro comparsa.

L'idea di analisi asintotica è quindi di "portare all'estremo" questo comportamento: se si è interessati a quei materiali dove i vortici appaiono – cioè per i quali questa costante è grande – perché non guardare ciò che accade quando la si rende *infinitamente grande* matematicamente? Questa idea risulta effettivamente essere molto fruttuosa nell'analisi di equazioni o di energie provenienti da modelli fisici. La filosofia, che è molto generale, è di arrivare a ridurre un complicato problema di minimizzazione (in uno spazio ad infinite dimensioni) ad un problema più semplice, con un minor numero di dimensioni (vale a dire, con meno parametri possibili da considerare). La difficoltà è quella di giustificare rigorosamente questa semplificazione (in qualche modo, di dimostrare che è "legale") e dare la formula esplicita per l'energia ridotta.

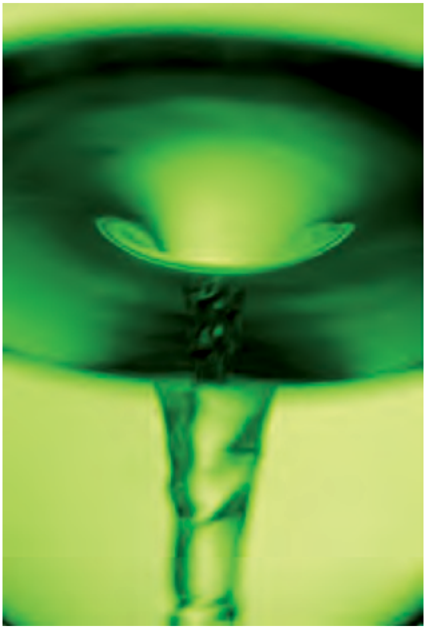
La miglior disposizione possibile

Solo molto di recente siamo riusciti ad esplicitare, dalla Ginzburg-Landau, la formula dell'energia semplificata, che disciplina la distribuzione e la ripartizione dei vortici. Questa energia non è sconosciuta né ai fisici, né ai matematici: essa corrisponde ad una interazione di tipo elettrostatico tra i vortici. Essi pertanto si comportano come delle cariche elettriche puntiformi che si respingono tra loro, ma che sono limitate allo stesso tempo da una forza esterna (la forza del campo magnetico applicato). È la competizione tra questi due effetti a dettare la loro distribuzione. Per fare un'analogia, immaginiamo molte persone che si detestano, ma sono costrette a stare tutte insieme nella stessa stanza: qual è il modo migliore di posizionarsi per rimanere più lontani possibile gli uni dagli altri?

Questo problema non si limita alla superconduttività: è in effetti un tipo di domanda comune in natura. Perché gli atomi in un cristallo si dispongono proprio secondo un reticolo cristallino (in generale cubico), vale a dire in modo periodico? Di nuovo, è la somma di tutte le interazioni elettrostatiche tra gli atomi, e tra gli atomi e gli elettroni, a dettarne la disposizione. Possiamo misurare il costo di ogni distribuzione tramite una energia, per la quale si suppone che il minimo venga proprio raggiunto da una tale disposizione in rete. Ahimè, non si sa dimostrare che queste strutture cristalline sono quelle che minimizzano questa energia, anche se le osserviamo chiaramente in natura ... Il solo caso in cui sappiamo dimostrare qualcosa di simile è quando cerchiamo la migliore distribuzione di palline rigide identiche in un piano, come le palle da biliardo, quando se ne vuole mettere il più possibile in una unità di volume. In questo caso, è stato dimostrato assai di recente che la migliore configurazione è un reticolo esagonale (in 3 dimensioni, allo stesso modo in cui le arance vengono impilate sul banco del fruttivendolo!)

Questioni di periodicità come queste sono quindi generali, fondamentali in natura, e costituiscono una vera e propria sfida matematica. Nel caso della repulsione elettrostatica che governa l'interazione tra vortici in Ginzburg-Landau, si pensa che la soluzione migliore sia di nuovo il reticolo esagonale, di nuovo la famosa rete di Abrikosov! Ma per il momento, non si riesce a dimostrarlo rigorosamente. La Ginzburg-Landau non ha quindi ancora rivelato tutti i suoi segreti ...

La migliore configurazione per delle sfere rigide identiche in un piano, come le palle da biliardo, è la rete esagonale (in dimensione 3, allo stesso modo in cui le arance vengono impilate sul banco di un fruttivendolo!)



I(n)spirazione matematica:
la modellizzazione del polmone

Céline Grandmont, *direttrice di ricerca presso Inria*

Il nostro sistema respiratorio, per la sua complessità, è un eccellente oggetto per l'indagine matematica. Il funzionamento del sistema respiratorio è descritto da equazioni che vengono utilizzate per eseguire simulazioni che completano le sperimentazioni e permettono di comprendere meglio e di prevedere i fenomeni che si verificano quando respiriamo.

Quali sono i fenomeni fisiologici che si verificano durante un attacco d'asma? Dove si depositano nei polmoni gli aerosol terapeutici o le particelle inquinanti emesse dai gas di scarico? Qual è il modo migliore per ventilare un paziente che ha bisogno di assistenza respiratoria? In che modo enfisema o fibrosi, malattie del tessuto polmonare, alterano la ventilazione e la diffusione dell'ossigeno nel sangue?

Tutte queste domande interessano i medici – pneumologi, allergologi – i biologi, gli studiosi di meccanica dei fluidi, ma anche i matematici, che cercano di fornire alcune risposte.

L'albero bronchiale: una geometria complessa

La respirazione coinvolge molti fenomeni fisici, in scale molto diverse. La funzione principale del sistema respiratorio è quella di garantire lo scambio di gas tra aria e sangue. L'aria inspirata attraverso il naso, o la bocca, viene poi trasportata nelle vie respiratorie: faringe, laringe, trachea, bronchi ... L'albero bronchiale umano possiede una geometria molto complessa: si presenta in forma di un albero *diadico* (in cui ogni ramo si divide in due e così di seguito) con circa 23 livelli di biforcazione e ogni generazione è costituita da condotti più piccoli dalla precedente, con lunghezze che diminuiscono da 10 cm (trachea) a 1 mm (gli ultimi bronchioli respiratori). Questo albero è circondato da un materiale viscoelastico, il *parenchima polmonare*, consistente tra l'altro di elastomero, di fibre di collagene e di una rete di vasi sanguigni.

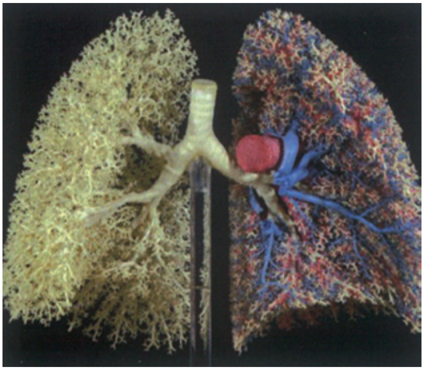


Figura 1. Modello di polmone umano. Istituto d'Anatomia dell'Università di Berna, Prof. Ewald R. Weibel.

La prima parte dell'albero bronchiale, fino alla quinta generazione, è puramente conduttrice e ha per funzione principale il trasporto dell'aria. Successivamente, degli alveoli sono innestati sui bronchioli e consentono lo scambio di gas (ossigeno e anidride carbonica) con i capillari sanguigni, i quali foderano l'esterno con circa 300 milioni di alveoli. Il movimento dell'aria in questa rete di tubi è assicurato dal diaframma e dalla gabbia toracica, i veri motori della respirazione. Questo si traduce, durante l'ispirazione, in aumento del volume polmonare, il quale induce una differenza di pressione tra la bocca (o il naso) e gli alveoli polmonari, creando un flusso d'aria nell'albero (da questo punto di vista, il polmone assomiglia a un mantice).

Non sappiamo necessariamente misurare sperimentalmente tutte le quantità desiderate e non possiamo ripetere gli esperimenti all'infinito per confutare o confermare certe ipotesi. Per questo motivo può essere utile fare un modello.

Descrivere e comprendere il trasporto di aria, o il deposito di particelle, in questa geometria complessa, e in tutta la sua generalità, può rivelarsi difficile. Non sappiamo necessariamente misurare sperimentalmente tutte le grandezze desiderate e non possiamo ripetere gli esperimenti all'infinito per confutare o confermare certe ipotesi. Questo è il motivo per cui può essere utile modellizzare – vale a dire, in questo contesto, descrivere con equazioni e effettuare delle simulazioni al computer – in un modo semplice, ma rappresentativo, eventi specifici, come la variazione di volume polmonare globale durante i cicli della respirazione. Si può anche cercare di descrivere con maggiore dettaglio una particolare area, ad esempio la porzione superiore (detta *prossimale*) dell'albero, al fine di ottenere una mappatura del flusso d'aria in questa regione.

La modellizzazione, naturalmente, dipenderà dalle necessità che la motivano. Potrebbe trattarsi di capire un meccanismo globale e l'effetto che certe patologie hanno su questo meccanismo, oppure per ottenere una descrizione tridimensionale del flusso, ad esempio per prevedere in quali aree dell'albero si depositano delle particelle curative o inquinanti.

Un modello semplificato ...

Riprendiamo l'immagine del mantice. Immaginiamo di cercare di riprodurre i risultati di un esperimento di *spirometria* (dal latino *spiro*, respirare, e dal greco *metron*, misura: misurazione della respirazione). In questo esperimento, al paziente viene chiesto di inspirare ed espirare più forte possibile e viene misurato il volume respirato, che indicheremo con V, e il flusso associato (che non è altro che il tasso di variazione di volume nel tempo), che indicheremo con $\dot{V}$.

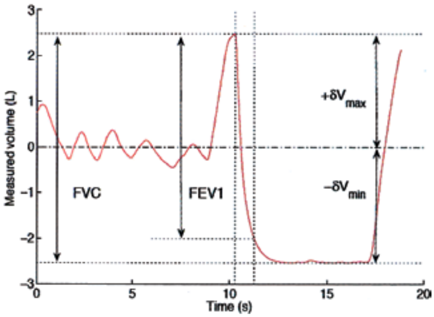


Figura 2. Evoluzione del volume di respirazione nel tempo (da S. Martin, T. Similowski, C. Straus, B. Maury, ESAIM Proc., Vol. 23, 2008, p. 30-47)

Supponiamo di voler descrivere il comportamento globale della dinamica dei volumi di respirazione (vale a dire come cambia questo volume respirato nel tempo). Il modello matematico più semplice, ampiamente utilizzato, che possiamo applicare è quello che è chiamato il *modello monocompartimento*. In questo modello, noto ed utilizzato da lungo tempo, il flusso d'aria è proporzionale alla differenza tra la pressione alveolare P_a e quella alla bocca P_{atm}. La costante di proporzionalità è la resistenza R dell'albero bronchiale al flusso (la resistenza dipende dalle dimensioni del tubo e dalla viscosità del fluido, in particolare, più il tubo è lungo e sottile e più è difficile far passare l'aria all'interno). Questa può essere espressa da un'equazione chiamata Legge di Poiseuille dal nome del medico francese che l'ha stabilita nel 1844, nella quale si afferma:

$$P_{atm} - P_a = R\dot{V}$$

Inoltre il volume del "palloncino" dipende dall'elastanza dei tessuti (indicata con E) e dalle forze ad esso applicate (nel nostro caso dalla pressione alveolare P_a e dalla pressione P esercitata dalla membrana). Questo è espresso da un'altra equazione che traduce l'equilibrio elastico del palloncino:

$$P_a - P = EV$$

Infine, combinando le due equazioni, si ottiene:

$$P_{atm} - P = R\dot{V} + EV$$

che descrive l'evoluzione nel tempo del volume del palloncino.

Si tratta di un'equazione differenziale, poiché fa intervenire sia il volume V che la sua variazione $\dot{V}$, lineare, per la quale sappiamo calcolare esplicitamente la soluzione.

Per sfruttare tale modello, per quanto semplice sia, bisogna trovare i valori dei parametri che sono la resistenza delle vie aeree, l'elastanza dei tessuti e la pressione esercitata.

Questa ricerca dei giusti parametri può essere fatta "a mano" o, se siamo in grado di misurare la pressione esercitata, in modo sistematico mediante metodi matematici al fine di produrre la soluzione più precisa del volume misurato.

Aggiustarli a "mano" significa provare diversi valori dei parametri per riprodurre le migliori misure di flusso di volume, in accordo con i dati sperimentali. Se la pressione esercitata è misurabile, è anche possibile risolvere un problema di minimizzazione, dove si cercano R ed E che minimizzano la differenza tra la pressione misurata P_m e la pressione P data dall'equazione:

$$P = P_{atm} - R\dot{V} - EV$$

dove V e $\dot{V}$ sono dati sperimentalmente.

Così, una volta "aggiustati" i parametri, risolvendo la nostra equazione differenziale, si ottiene una formula matematica che esprime il volume in funzione del tempo e di P_{atm}, P_a, P e E:

$$V(t) = V(0)e^{-\frac{E}{R}t} + \int_0^t \frac{P_{atm} - P(s)}{R} e^{\frac{E}{R}(s-t)} ds$$

... che può essere arricchito

Anche se si è interessati soltanto ad una dinamica globale del volume, questo modello semplificato non tiene conto di tutti i fenomeni che possono avere un impatto su ciò che si sta cercando di misurare. In particolare, l'espansione polmonare è limitata dalla cassa toracica, l'albero bronchiale si deforma durante i cicli di respirazione, la resistenza al flusso aumenta con la velocità... È quindi necessario arricchire il modello.

Si può immaginare che la resistenza R e l'elasticità E dipendano dal volume V e dal flusso $\dot{V}$. Ma tenendo conto di questa dipendenza, giungiamo a una equazione per la quale non sappiamo più esplicitamente calcolare la soluzione. Tuttavia, siamo in grado di dire se una tale soluzione esiste e di fornire dei metodi per calcolarne un'approssimazione. Il lavoro del matematico applicato è quindi quello di comprendere il comportamento delle soluzioni e calcolare, al variare del tempo e con l'ausilio del computer, il volume di respirazione.

Se la soluzione del modello matematico e numerico è vicina alla sperimentazione spirometrica, allora il modello è "buono": si può in seguito "giocare" con i parametri e farli variare per riprodurre o per capire i fenomeni in gioco. La resistenza R aumenta durante un attacco d'asma, l'elastanza E del tessuto diminuisce in caso di enfisema. Possiamo quindi contribuire ad una prospettiva diversa da quella delle misure sperimentali ed aiutare a rispondere a domande quali:

In che modo un enfisema o una crisi d'asma influenzano globalmente la ventilazione?

Ovviamente, un tale modello non può spiegare il flusso di aria in tutta la sua complessità e non permette di conoscere la velocità dell'aria nella trachea o nei piccoli bronchi. Ciò non di meno, se cerchiamo di descrivere il comportamento dell'aria in ciascun punto dell'albero bronchiale, l'approccio del matematico applicato sarà lo stesso: determinare un sistema di equazioni capace di spiegare il più fedelmente possibile la realtà, capire tali equazioni e le proprietà delle loro soluzioni utilizzando strumenti matematici, trovare strategie – dal momento che non sappiamo calcolare le soluzione esatte – per determinare approssimazioni numeriche, utilizzare questi calcoli e confrontarli con le misure sperimentali disponibili per convalidare o invalidare il modello ...

Con diversi livelli di descrizione

Supponiamo dunque che si cerchi di descrivere il flusso d'aria nell'albero bronchiale, non più unicamente mediante il volume respiratorio, ma in modo più accurato. Ad esempio, determiniamo la velocità e la pressione del fluido in ogni punto di quest'albero durante un ciclo respiratorio. Data la complessità geometrica, i calcoli non sono, allo stato attuale, improponibili sulla totalità dell'albero. Possiamo quindi, in un primo tempo, limitarci alla parte superiore dell'albero respiratorio (fino alla decima generazione). Chiamiamo questa zona Ω.

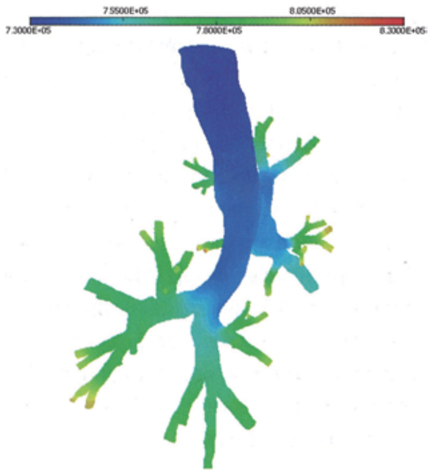


Figura 3. Zona Ω.

All'interno di Ω, è ragionevole supporre che la velocità e la pressione dell'aria verifichino delle equazioni ben note nella meccanica dei fluidi: le equazioni di Navier-Stokes (vedi riquadro p. 49). Tuttavia, questa velocità e questa pressione sono altamente dipendenti da ciò che accade oltre la decima generazione e, in particolare, dal movimento del diaframma e dell'elasticità del tessuto polmonare. Dobbiamo trovare un modo per descrivere matematicamente questa parte del sistema respiratorio e accoppiarlo con le equazioni di Navier-Stokes. Possiamo ispirarci al semplice modello descritto in precedenza. Otteniamo in questo modo un sistema di equazioni con due livelli di descrizione: una descrizione dettagliata di Ω e una più grossolana per le zone restanti. Il passo successivo è quello di studiare il sistema risultante: esiste una soluzione e quali sono le sue proprietà matematiche? Come calcolare efficacemente le velocità e le pressioni approssimate? Questo sistema rappresenta in parte la realtà?

Le risposte alla prima domanda permettono di fornire un quadro di riferimento per comprendere meglio il dominio di validità del modello. Le difficoltà qui sono sono legate al fatto che siamo in presenza di equazioni *non lineari*, che mettono insieme diversi livelli di descrizione. Inoltre, guardiamo il flusso in una data zona e "tagliamo" artificialmente ciò che non ci interessa. Quindi si tratta di un sistema aperto in cui la massa e l'energia (cinetica) possono variare. Uno dei problemi matematicamente importanti con il quale dobbiamo confrontarci è quindi di stimare questa energia.

Tutti questi passaggi pongono domande appassionanti che hanno un interesse matematico proprio.

Questo problema è cruciale anche quando si cerca di rispondere alla seconda domanda, sul calcolo numerico della velocità e della pressione. Questo è un calcolo approssimato, perché, come abbiamo detto, non è noto come produrre soluzioni esplicite (cioè scrivere la velocità in ogni punto e in qualsiasi momento, come per il volume nella formula di p. 46). Per questo, cercheremo di descrivere il campo della computazione Ω, che è ottenuto a partire dalle immagini mediche ed è unico per ogni paziente, con un numero finito di punti: *nodi* connessi da *archi*. È quello che viene chiamato un reticolo. Si prende un'immagine da una risonanza magnetica o da una TAC e la si "segmenta" e la si riduce ad un reticolo: questo, di nuovo, richiede significativi sviluppi matematici, necessari per queste geometrie complesse, costituite da curve con diversi rami come quelle dell'albero bronchiale.

Figura 4. Magliatura di Ω.

È nei suddetti nodi che verrà determinata la velocità e la pressione del fluido.

Infine, calcolare numericamente i valori approssimati delle velocità e delle pressioni in questi punti dello spazio e in un numero finito di istanti, richiede l'individuazione di metodi specifici adattati al modello che, ad esempio, trattino efficacemente il fatto che il sistema è "aperto".

Tutti questi passaggi pongono appassionanti domande che hanno un interesse matematico proprio e che sono attualmente oggetto di filoni attivi di ricerca. Malgrado tutto, non è possibile realizzare questi modelli senza confrontare i risultati della simulazione con gli esperimenti. I matematici interagiscono con gli specialisti delle scienze respiratorie, che possono confermare se i modelli sviluppati sono pertinenti e come questi possono essere migliorati. L'obiettivo finale è quello di ottenere strumenti di previsione affidabile e fornire un chiarimento complementare a quello che si ottiene attraverso la sperimentazione *dal vivo o in vitro*.

Le equazioni di Navier-Stokes

Le equazioni di Navier-Stokes qui utilizzate descrivono il flusso di un fluido viscoso incomprimibile e quindi collegano la velocità u del fluido con la sua densità ρ e la pressione p. Esse si scrivono

$$\begin{cases} \rho\partial_t u + \rho(u \cdot \nabla)u - \nu\Delta u + \nabla p = 0 \\ \nabla \cdot u = 0 \end{cases}$$

La prima equazione esprime il principio fondamentale della dinamica di Newton in assenza di forze esterne, mentre la seconda traduce l'incomprimibilità del liquido, la cui viscosità è indicata da ν. Per risolvere queste equazioni numericamente, è necessario fornire la velocità al bordo della zona in cui sono applicate, oltre che in tutto il dominio all'istante iniziale.

Il tempo che farà

Basdevant Claude, *professore presso l'Università di Parigi 13 e presso École Polytechnique*

Le previsioni meteorologiche o climatiche non sono un'impresa facile. Esse implicano la modellizzazione di numerosi fenomeni di diversa natura e l'intervento di diverse scienze, dalla matematica alla biologia, passando per l'informatica, la fisica o la chimica.

Dietro le piccole nuvole grigie e i soli radiosi che punteggiano la mappa nel bollettino meteorologico della sera, è da tempo che non ci sono più rane[1] e termometri, ma computer super-potenti ai quali abbiamo fatto assorbire un gran numero di misurazioni (ottenute principalmente via satellite), molte leggi della meccanica e della fisica, ma anche tanta matematica, a volte molto recente.

Affinché i computer possano fornire le previsioni, è necessario sviluppare preliminarmente quello che viene chiamato un *modello numerico delle previsioni del tempo*. Schematicamente, un tale modello di previsione a scadenza da otto a dieci giorni rappresenta lo stato dell'atmosfera in ogni momento, attraverso i valori dei parametri meteorologici (velocità del vento, temperatura, umidità, pressione, nuvole, aerosol, ecc) al centro di "scatole" da due a venti chilometri di lato e da qualche decina a centinaia di metri d'altezza. Questo partizionamento immaginario dell'atmosfera in scatole è necessario perché è impossibile specificare i parametri in tutti i punti dell'atmosfera (questi punti sono un numero infinito!). In linea di principio, più le caselle sono piccole – e numerose –, più la descrizione dello stato atmosferico è preciso e più le previsioni lo saranno a loro volta. Ma in pratica, le scatole sono non meno di un qualche chilometro per una previsione regionale a breve termine, e qualche decina di

1 Secondo una credenza particolarmente diffusa in Francia si può prevedere il tempo osservando il comportamento delle rane, N.d.T.

chilometri per una previsione a dieci giorni che è necessariamente relativa a tutto il pianeta; oltre questo, neanche la potenza dei computer più grandi è sufficiente: è necessario che il calcolo venga completato in tempo, vale a dire, in molto meno di 24 ore!

Partendo dallo stato dell'atmosfera all'inizio del periodo da prevedere, il modello fa calcolare al computer la sua futura evoluzione usando le leggi della dinamica e della fisica. L'evoluzione nel tempo viene calcolata passo a passo, ad intervalli di pochi minuti. Questo è il principio di previsione del tempo numerica, un principio noto dall'inizio del ventesimo secolo, ma che ha dovuto attendere gli anni 1940-1950 e i primi computer prima di essere messo in atto.

Misurazioni inutilizzabili direttamente

Il primo problema nello schema di previsione ideale che è stato appena descritto è costruire lo *stato iniziale dell'atmosfera*. Le osservazioni sono lungi dall'essere adatte a questo scopo. Le stazioni meteorologiche sono irregolarmente distribuite sul globo e forniscono pochissime misure in altitudine. Quanto ai satelliti, essi sono per la maggior parte a scorrimento, vale a dire, scansionano in continuo la Terra. Le loro misurazioni non sono, dunque, ottenute al medesimo istante in tutti i punti. Inoltre, i satelliti misurano delle quantità integrate su tutto lo spessore dell'atmosfera (si tratta solitamente del flusso di energia ricevuta in un certo intervallo di lunghezze d'onda) e non le variabili meteorologiche (vento, temperatura, umidità, ecc.) che sono quelle che entrano in gioco nelle equazioni dei modelli.

Abbiamo, quindi, una massa di dati disparati, mal distribuiti sulla superficie del globo, che si estendono su 24 ore, acquisiti con approssimazioni variabili, con i quali si deve *inizializzare* una previsione, vale a dire, costruire uno stato meteorologico iniziale rispetto al quale il modello simulerà l'evoluzione. Ma grazie al lavoro nel campo dell'*ottimizzazione dinamica*, nel quale hanno molto contribuito lo scienziato russo Lev Pontryagin (1908-1988) e la scuola matematica francese, sono stati messi a punto negli anni 1980, dei metodi detti di *assimilazione variazionale* che permettono di ricostruire in modo ottimale lo stato iniziale. L'idea alla base di tali metodi, operativi dagli anni 2000 presso Meteo-France, è quella di forzare in qualche modo la traiettoria del modello numerico, facendola passare "vicino" ai dati osservati durante le 24 ore precedenti, tuttavia preservando l'equilibrio del modello. L'assimilazione variazionale non è, d'altra parte, l'unica tecnica matematica moderna che ha rivoluzionato il trattamento delle osservazioni: l'uso di reti neuromimetiche o delle ondine, inventate negli anni 1980, ha portato a guadagni notevoli in termini di efficienza, precisione e rapidità nell'elaborazione dei dati forniti dai satelliti.

> *Si dispone, quindi, di una massa di dati disparati, mal distribuiti sulla superficie del globo, che si estendono su 24 ore, acquisiti con diversi gradi di approssimazione.*

Quando l’analisi numerica entra in azione

Una volta conosciuto lo stato atmosferico iniziale necessario per il modello di previsione numerica, resta ancora da scrivere il software in grado di calcolare il tempo futuro a partire da questo stato iniziale e dalle leggi della fisica. Queste leggi si basano su una descrizione continua dello spazio e del tempo. Invece il nostro modello numerico prende in considerazione soltanto un numero finito, seppur grande, di scatole. Inoltre, gli intervalli di tempo tra due stati calcolati sono di vari minuti - si dice che il problema è stato *discretizzato*. Passare da equazioni continue a degli schemi numerici per il modello discretizzato, mantenendo la migliore precisione possibile, è precisamente il campo dell’analisi numerica, una branca della matematica che è esplosa dopo l’arrivo dei computer. L’analisi numerica ha come obiettivo la risoluzione delle equazioni in modo esplicito, vale a dire, fino ad ottenere valori numerici accurati, investendo il minor tempo e sforzo possibile. Questo è indispensabile perché la simulazione sia utile e per valutare il grado di incertezza delle previsioni. Ad esempio, notevoli progressi sono stati recentemente compiuti sui metodi per simulare il movimento delle sostanze chimiche o delle particelle nella turbolenza atmosferica. Questi progressi hanno migliorato in modo significativo lo studio e la previsione dell’inquinamento atmosferico.

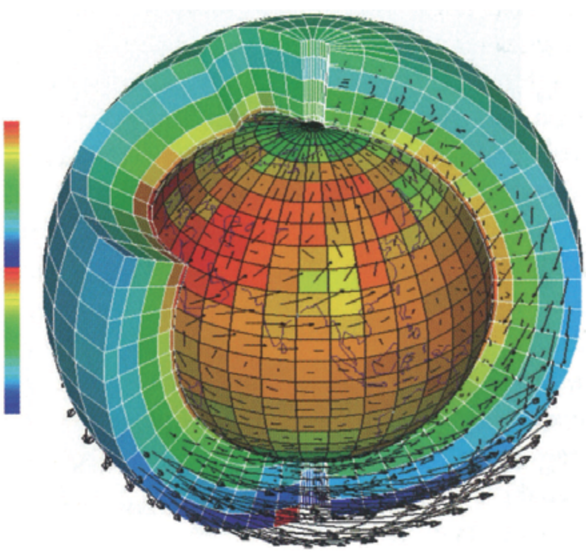


Figura 1. Vista prospettica delle scatole di un modello di previsione del tempo o del clima. (L. Fairhead LMD / CNRS)

Le previsioni meteo sono affidabili?

Una delle più grandi sfide che riguardano le previsioni meteo è quella di riuscire a valutare la qualità delle previsioni effettuate: si possono considerare affidabili le previsioni per tre, quattro o anche dieci giorni? Oltre che ad interessare chi cerca di organizzare il proprio lavoro o le proprie vacanze, questa domanda ha un forte impatto umano ed economico: sia per la previsione di eventi estremi che possono essere devastanti, che per la gestione dell'energia o per la produzione del gelato! Sebbene i progressi constatati a posteriori siano molto incoraggianti, rimane difficile stimare a priori l'affidabilità delle previsioni. Negli ultimi anni, alcuni strumenti nati dall'incrocio tra la teoria dei sistemi dinamici e la teoria della probabilità sono stati utilizzati per sviluppare quelle che vengono chiamate le *previsioni d'insieme*: questo consiste nel fare non più una sola previsione da un unico stato iniziale, ma effettuare in parallelo un gran numero di previsioni (ad esempio 50) a partire da condizioni iniziali perturbate. La teoria dei sistemi dinamici consente di scegliere le perturbazioni iniziali più significative, mentre la teoria della probabilità permette di estrarre dalla divergenza delle simulazioni delle informazioni sull'affidabilità di quest'ultime e di calcolare la probabilità che un evento specifico si verifichi.

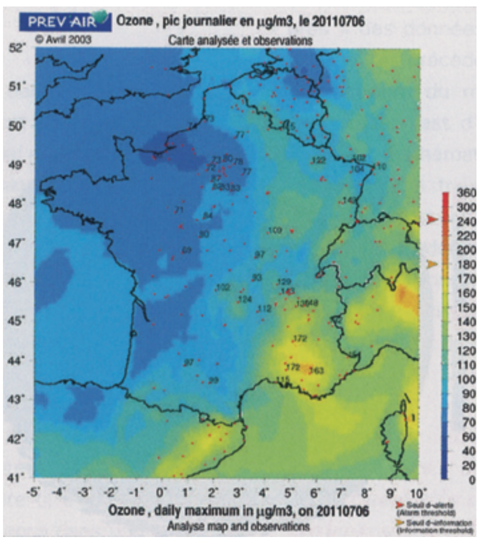


Figura 2. Le concentrazioni di ozono in superficie del 6 luglio 2011 rilasciato dal sistema di previsione nazionale della qualità dell'aria PREVAIR (INERIS / IPSL / Météo-France). I colori rappresentano le concentrazioni previste dal modello di chimica e trasporto CHIMERE, i valori numerici sono i picchi di concentrazione rilevati nella giornata dalla rete di monitoraggio.

Possiamo prevedere il tempo con largo anticipo? La teoria dei sistemi dinamici dice di no.

Abbiamo parlato finora delle previsioni a breve tempo, da otto a dieci giorni. Ma perché non facciamo delle previsioni per periodi più lunghi? Il meteorologo americano Edward N. Lorenz ha dimostrato in un famoso articolo del 1963, che, con tutta probabilità, non abbiamo speranza di riuscirci. L'atmosfera è un sistema caotico, vale a dire che qualsiasi errore sullo stato-tempo iniziale, per quanto possa essere piccolo, viene rapidamente amplificato nel tempo, talmente rapidamente che una previsione a scadenza di dieci giorni perde completamente senso. Questo non significa, tuttavia, che non si possa prevedere il clima - vale a dire fare una previsione di tipo statistico e non deterministico, studiare le medie delle temperature o delle precipitazioni in un periodo o in una regione, anziché il tempo specifico che farà a Quimper in Bretagna in un determinato giorno del mese di luglio. La questione è di grande importanza: il nostro clima futuro è minacciato dalle emissioni di gas causate dalle attività umane e dobbiamo prevedere l'effetto a lungo termine di questi mutamenti, conseguenza dell'effetto serra. La teoria dei sistemi dinamici ci fornisce gli strumenti per dare validità a questa modellizzazione del clima. Questo dominio, nel quale il matematico Henri Poincaré, all'inizio del XX secolo, è stato un grande precursore, ha registrato notevoli progressi negli ultimi anni del XX secolo. La teoria dei sistemi dinamici permette, ad esempio, di identificare ciò che i matematici chiamano *attrattori* o *modelli climatici* per i meteorologi. Questo permette anche di riconoscere quali sono i modelli climatici più prevedibili e quelli che sono più instabili. In situazioni di instabilità, uno strumento utilizzabile è la modellazione probabilistica del clima, vale a dire, la progettazione di modelli che tengano conto esplicitamente del carattere aleatorio della previsione. Ancora poco sviluppati, i modelli di questo tipo necessitano di strumenti molto recenti della teoria delle equazioni alle derivate parziali stocastiche.

Dalle previsioni meteorologiche alle previsioni climatiche

I modelli numerici per le previsioni del clima assomigliano ai modelli per le previsioni meteorologiche, con due differenze essenziali. Per ragioni dovute al tempo di calcolo, le loro "scatole" sono necessariamente più grandi (da 100 a 300 km quadrati); dato che l'intervallo di tempo da simulare va da pochi mesi a centinaia o addirittura migliaia di anni, è impossibile essere più precisi. Ma la differenza importante è che i cambiamenti climatici avvengono su scale temporali lunghe e che questo non permette più di trascurare le interazioni che intercorrono tra l'atmosfera, l'oceano, i ghiacci del mare e la biosfera. Questo è il motivo per cui un modello climatico deve combinare un modello per l'atmosfera, un modello per l'oceano, un modello per il ghiaccio dei mari, un modello per la biosfera e per i cicli degli scambi di energia, di quantità di moto, d'umidità, di gas, così come delle interazioni chimiche tra questi mezzi. Al di là della complessità informatica di una tale costruzione, sorgono dei delicati problemi matematici riguardo a quale sia il modo giusto di mettere in relazione questi domini e come specificare le condizioni nell'interfaccia atmosfera-oceano, oceano-ghiaccio, ecc. E affinché il calcolo nelle "grandi scatole" resti significativo, bisogna considerare l'effetto statistico, sulla scala di queste scatole, dei processi che si verificano su scale molto più piccole (ad esempio: qual è l'effetto statistico sul bilancio energetico di una scatola di 300 km dovuto alla presenza di piccoli cumuli di pochi km di diametro?). Restano ancora, in tutte queste questioni, molti aspetti interdisciplinari da sviluppare, per i quali la matematica detiene un ruolo importante.

Bibliografia

Jouzel J., Debroise A., (2004). *Le climat: jeu dangereux, quelques prévisions pour les siècles à venir*, éd. Dunod.

Le Treut H., Jancovici J.-M., (2004). *L'effet de serre: allons-nous changer le climat?*, éd. Flammarion.

Sadourny R., (2002). *Le climat est-il devenu fou?*, éd. Le Pommier.
- –, (2003). *Peut-on croire à la météo?*, éd. Le Pommier.
- –, (2005). *D'où viennent les tempêtes?*, éd. Le Pommier

Temam R., Wang S., (2000). *Mathematical Problems in Meteorology and Oceanography*, Bull. Amer. Meteor. Soc., 81, p. 319-321.

Ulteriori riferimenti:

Joussaume S., (2000). *Le Climat: d'hier à aujourd'hui*, CNRS Editions.

Hauglustaine D., Jouzel J., Le Treut H., (2004). *Climat: chronique d'un bouleversement annoncé*, éd. Le Pommier.

Internet, incendi boschivi e porosità: alla ricerca del punto in comune

Théret Marie, *docente presso l'Università Paris Diderot*

Lo scambio di dati tra gli utenti, la diffusione di un incendio boschivo, le infiltrazioni d'acqua in una roccia: un semplice modello matematico che utilizza i grafi permette di capire meglio questi fenomeni.

Che cosa c'è in comune tra una foresta, una rete di comunicazione e una roccia porosa? Per preservare le foreste è importante capire come si propaga un fuoco, o una certa malattia, da un albero all'altro. Per migliorare le reti di comunicazione è necessario sapere come i dati circolano tra i diversi utenti. Per studiare la porosità di una roccia bisogna descrivere come l'acqua filtra al suo interno. Questi tre problemi molto concreti, se pure hanno formulazioni molto diverse nel linguaggio corrente, possono essere rappresentati dallo stesso modello matematico: il modello della percolazione.

La diffusione di una malattia in un bosco

Concentriamoci sull'esempio del bosco per descrivere questo modello. Vogliamo studiare come si diffonde una malattia: la contaminazione causata da un albero malato rimarrà confinata in una piccola area, o riuscirà a diffondersi ad alberi che si trovano molto lontano rispetto alla zona di contaminazione iniziale? L'approccio che adottiamo è quello della fisica statistica: per capire un fenomeno macroscopico, vale a dire su larga scala, si studiano gli elementi microscopici che compongono il sistema. Il sistema studiato è l'intero bosco, costituito da un gran numero di elementi "microscopici" (vale a dire, molto piccoli rispetto al bosco): gli alberi.

Dobbiamo cominciare descrivendo come avviene il contagio da un albero all'altro. Si presume che la contaminazione diretta possa avvenire soltanto tra alberi sufficientemente vicini, e che essa si verifichi con una certa probabilità, che dipende solo dalla virulenza della malattia (e non dalle caratteristiche individuali degli alberi, o dalla zona dove sono). Queste ipotesi sono da un lato entrambe credibili dal punto di vista biologico, dall'altro costituiscono una semplificazione sufficiente da permettere un trattamento matematico efficiente.

Essi formano un grafo in cui ogni albero è un vertice e dove si collegano due alberi da un arco se sono sufficientemente vicini da potersi contaminare uno con l'altro.

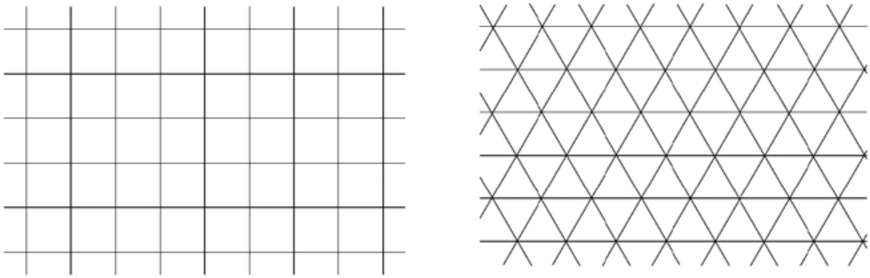


Figura 1. Due grafi ampiamente usati in teoria della percolazione: il reticolo quadrato e il reticolo triangolare

Il grafo risultante descrive come la malattia potrebbe diffondersi nello scenario peggiore, ovvero se ogni albero infettasse tutti i suoi vicini. Ma vogliamo considerare il fatto che la trasmissione della malattia si verifica in modo aleatorio. Pertanto, vi è una certa probabilità di contaminazione tra vicini, quindi coloreremo un arco con un pennarello rosso quando si verifica la contaminazione lungo un arco.

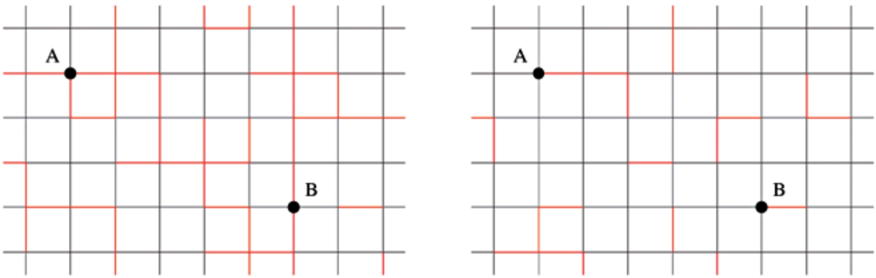


Figura 2. Archi colorati in modo casuale, indipendentemente l'uno dall'altro, ad indicare che la contaminazione tra due alberi adiacenti si verifica. Questa contaminazione si presume istantanea. In particolare, l'ordine in cui coloriamo gli archi non influisce. A sinistra: L'albero A contamina l'albero B e anche altri alberi più lontani. Questo avviene tipicamente quando la probabilità di trasmissione è alta. A destra: Un albero contamina soltanto altri 3 altri alberi. Non contamina B. Questo si verifica in genere quando la probabilità di trasmissione è bassa.

Quello appena descritto è il modello di percolazione ad archi: partendo da un dato grafo, abbiamo colorato in modo aleatorio gli archi secondo una certa probabilità, indipendentemente gli uni dagli altri.

Osserviamo che questo modello semplificato non tiene conto dell'aspetto temporale della diffusione della malattia: tutte le contaminazioni tra alberi vicini si verificano simultaneamente, istantaneamente e indipendentemente le une dalle altre. Gli archi non colorati corrispondono a coppie di alberi tra cui la malattia non si è trasmessa.

Quello che abbiamo appena descritto è il modello di percolazione ad archi. Per tradurre la differenza di scala tra l'albero e la foresta, si può considerare che vi è un numero infinito di alberi nella foresta, vale a dire che il grafo da studiare è infinito. Le domande che ci siamo posti sulla trasmissione della malattia possono essere riformulate in questo contesto: per sapere se l'epidemia causata da un albero malato rimarrà circoscritta, o se, al contrario, essa rischia di raggiungere altri alberi molto distanti, bisogna guardare quali vertici sono collegati tra loro mediante un cammino rosso al punto corrispondente all'albero malato iniziale. In particolare, ci chiediamo se questi alberi contaminati sono in numero finito o infinito. Il nostro modello dipende da un singolo parametro: la probabilità di contaminazione tra vicini, e si tratta dunque di studiare, a seconda dei valori di questo parametro, le proprietà del grafo aleatorio sopra descritto.

Questo modello è adeguato anche per descrivere una rete di comunicazione o una roccia porosa. Nel primo caso, gli archi del grafo sono i collegamenti della rete che possono essere in perfetto stato di funzionamento (arco colorato) oppure guasti o insufficienti (arco non colorato): possiamo in questo modo studiare la trasmissione delle informazioni inviate da un utente. Nel secondo caso, gli archi del grafo sono dei tubi microscopici nella roccia, che possono lasciare filtrare l'acqua, ma anche intasarsi. Ci si interessa, quindi, alla zona della roccia che è bagnata per capire se l'acqua filtra in un'area particolare. Il termine "percolare" significa, per un liquido, passare attraverso materiali porosi: è quindi questa interpretazione della roccia porosa che ha dato il nome al modello matematico della percolazione, ed è infatti per modellizzare i mezzi porosi che questo modello è stato introdotto da Broadbent e Hammersley nel 1957.

Un semplice ma ricco modello

Il modello della percolazione è molto semplice, ma anche molto ricco.
In primo luogo, perché presenta una transizione di fase, cioè un improvviso cambiamento di comportamento quando il parametro del sistema – la probabilità che un arco del grafo sia colorato – supera un certo valore critico. Infatti, se ogni arco è colorato con una probabilità sufficientemente elevata, la trasmissione della malattia si verifica in tutto il sistema, vale a dire, un albero cattivo può infettare altri alberi arbitrariamente lontani da lui. Matematicamente, ciò è espresso dicendo che "esiste una componente connessa infinita nel grafo aleatorio". Ma non appena la probabilità di trasmissione scende sotto una certa soglia, la trasmissione della malattia rimane locale, vale a dire che l'infezione resta limitata a piccole aree del bosco: "non esistono componenti connesse infinite" (vedi le figure nella pagina seguente). Questo teorema è stato dimostrato da Broadbent e Hammersley.

D'altra parte, nonostante la sua semplicità, questo modello solleva molte domande che rimangono senza risposta, a cominciare dalla più ovvia: come possiamo calcolare questo valore critico? Grandi progressi nello studio della percolazione si sono verificati in questi ultimi anni, in particolare attraverso l'utilizzo di strumenti dell'analisi complessa, un altro ramo della matematica (citiamo ad esempio i lavori di Lawler, Schramm e Werner degli anni 2000), ma la ricerca continua.

Nonostante la sua semplicità, questo modello solleva molte domande che rimangono senza risposta, a partire dalla più ovvia: come calcolare il valore critico?

Per uno studio più dettagliato di questi fenomeni, Hammersley e Welsh hanno introdotto nel 1965 un modello più complesso, il modello della percolazione di primo passaggio. Invece di colorare a caso degli archi del grafo, gli viene associato un tempo aleatorio (il tempo ri-

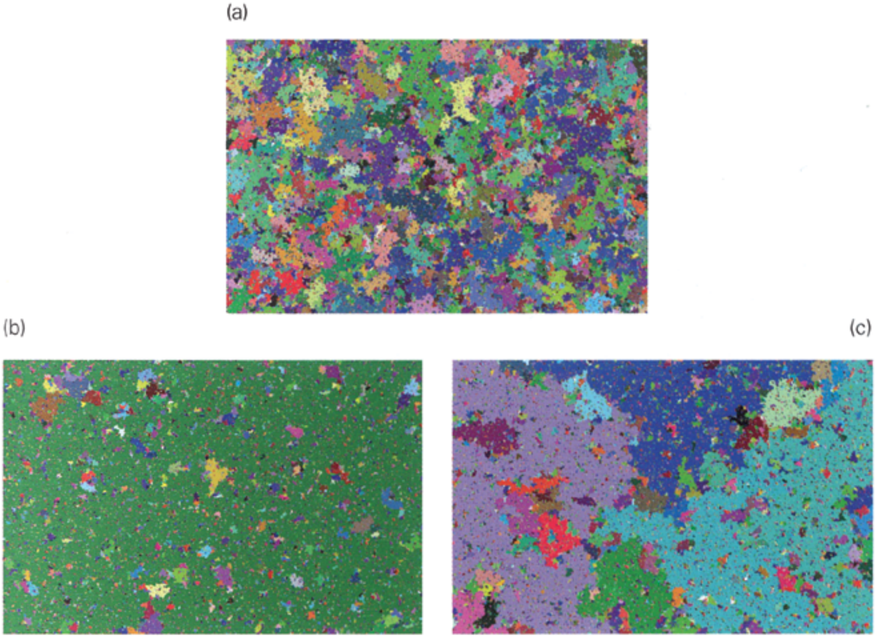


Fig 3. Fenomeno di transizione di fase (simulazioni al computer): analizziamo il sistema su larga scala, ogni pixel rappresenta qui un albero nella foresta. Un albero contamina quelli situati nella zona avente il suo stesso colore. (A) La probabilità di trasmissione è di 0,4. Le macchie di colore sono piccole: un albero malato infetta solo una regione circoscritta intorno ad esso. (B) p = 0,506, leggermente superiore al valore critico. La zona verde si estende indefinitamente: Un albero può contaminare alberi situati arbitrariamente lontani da lui. (C) p = 0.5, il regime critico, è la transizione tra gli altri due, il regime meno conosciuto. Le macchie sono di dimensione finita, ma molto grandi. Simulazioni numeriche di R. Cerf.

chiesto al fuoco di propagarsi tra due alberi o all'acqua per passare attraverso un tubo) o una capacità aleatoria (la massima quantità di informazioni che può passare attraverso un cavo in un secondo). Le questioni sollevate sono quindi più numerose: quanto tempo ci vuole perché l'acqua attraversi uno strato poroso di roccia? Quali saranno gli alberi colpiti da un incendio boschivo nelle prime ventiquattro ore? Qual è la quantità massima di informazioni che un utente della rete è in grado di inviare a un altro destinatario in un secondo? Molti teoremi sono stati ottenuti (citiamo i risultati negli anni '80 di Kesten, e più di recente i lavori di Benjamini, Kalai e Schramm nel 2003 e Chatterjee e Dey nel 2009), ma restano ancora molte sfide da affrontare!

Per andare oltre

Duminil-Copin H., *La Percolation, jeu de pavages aléatoires*, http://images.math.cnrs.fr/La-percolation-jeu-de-pavages.html

Grimmett G., (1989). *Percolation*. Springer-Verlag.

Alla ricerca della forma ideale

Grégoire Allaire, *professore presso École polytechnique*
François Jouve, *professore presso l'Università Paris Diderot*

Gli oggetti nati dalla produzione industriale sono concepiti in modo da ottimizzare una serie di parametri come il peso e la resistenza. Per evitare di cercare a tentoni la migliore forma possibile, possiamo oggigiorno contare su diversi metodi matematici di ottimizzazione.

La nostra società moderna è affascinata dal *design*, termine che riflette il nostro impegno a coniugare la bellezza all'utilità. Il grande pubblico conosce bene gli stilisti famosi e resi celebri dai media come Pininfarina o Starck, ma molto meno gli scienziati, gli ingegneri e i ricercatori, che si occupano di *design ottimale*: lontani da tutte le preoccupazioni estetiche, migliorano le forme degli oggetti industriali (struttura meccanica, profilo aerodinamico, elettronica, ecc.), al fine di aumentarne le prestazioni (resistenza, efficienza), pur tenendo conto dei vincoli, a volte contrapposti, come il peso o il costo. È evidente, ad esempio, che la resistenza di una struttura varia inversamente al suo peso (ciò che è più pesante è più robusto di ciò che è leggero). Così, l'ottimizzazione della robustezza di un aereo è limitata dal vincolo di un consumo minimo di carburante, che è direttamente collegato al peso. Un problema matematico classico è proprio quello di trovare la *soluzione ottimale* di un problema di ottimizzazione di una funzione (chiamata *funzione obiettivo*) nel rispetto dei vincoli.

> *È chiaro che la solidità delle strutture varia inversamente al suo peso (ciò che è pesante è più robusto di quello che è leggero).*

Il metodo tradizionale di ottimizzazione procede per tentativi ed errori, secondo la pratica e l'intuizione dell'ingegnere: si sceglie una forma della quale si calcolano le prestazioni e, a seconda del risultato, si apportano delle modifiche per cercare dei miglioramenti, fino all'ottenimento di una forma soddisfacente (in mancanza di poter trovare quella ottimale). Questo approccio "manuale" è molto lento, costoso e impreciso. Grazie all'enorme sviluppo della potenza di calcolo dei computer e al progresso matematico, questo approccio empirico viene progressivamente sostituito da applicativi software che automatizzano questo processo di ottimizzazione.

Ottimizzare la geometria con il metodo Hadamard

Qualsiasi algoritmo di ottimizzazione è *iterativo*: si costruisce una nuova forma a partire da una variazione della precedente. Successivamente si calcolano le prestazioni di questa nuova forma, le quali vengono confrontate con quelle della precedente. Infine, se la prestazione della struttura risulta migliorata, si ripete il procedimento a partire dalla nuova forma.

È necessario indovinare la topologia corretta da imporre alla forma di partenza, o questo sarà impossibile nella maggior parte dei casi.

Nel 1907, il matematico Jacques Hadamard ha proposto un metodo di variazione della forma, che oggi porta il suo nome, e che, nonostante la sua origine teorica, viene applicato in pratica per simulare alcuni problemi al computer. Il metodo consiste nel partire da una forma iniziale, spostando i bordi gradualmente, senza crearne nuovi. Questo metodo modifica la *geometria* della forma originale, ma ne conserva la *topologia*. Infatti le forme ottenute successivamente preservano lo stesso numero di

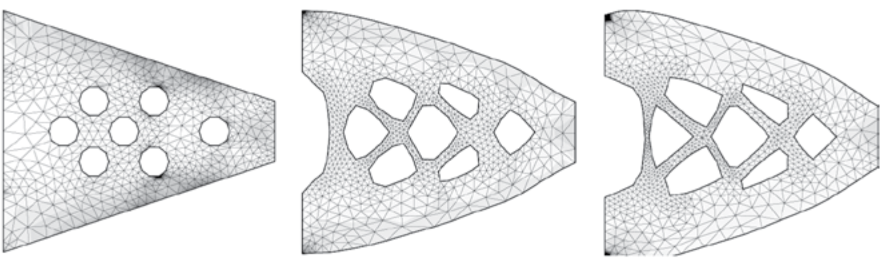


Figura 1. Configurazione iniziale (a sinistra), l'iterazione intermedia (al centro) e la forma ottimale (a destra) di una mensola, ottenute con il metodo di Hadamard.

fori, come mostrato nelle figure che rappresentano i risultati numerici del metodo Hadamard.

Il fatto che la topologia non cambi è una limitazione piuttosto fastidiosa. In realtà, ciò significa che è necessario indovinare la topologia giusta da imporre alla nostra forma di partenza, dal momento che non potremo cambiarla successivamente per migliorarne le prestazioni. Ma questo risulta essere impossibile nella maggior parte dei casi. Da qui la motivazione dei matematici ad inventare altri metodi in grado di ottimizzare la topologia, cioè il numero di fori.

Inoltre, il metodo di Hadamard presenta anche lo svantaggio di essere molto costoso in tempo di calcolo.

L'importanza dei materiali compositi

Nel 1990, i matematici hanno inventato un metodo di ottimizzazione della forma topologica, chiamato *metodo di omogeneizzazione*, che è oggi ampiamente utilizzato dagli ingegneri in molti software industriali.

L'idea di base è quella di trasformare un problema di ottimizzazione della forma in un problema di ottimizzazione della *densità di un materiale*.

In ogni punto dello spazio, la densità del materiale è un valore compreso tra 0 e 1. Il valore 0 corrisponde ad un foro o vuoto (assenza di materiale), il valore 1 corrisponde al materiale solido e i valori intermedi (ad esempio il valore 0,5) corrispondono ad un materiale composito poroso, ad esempio come una spugna. Più il valore è prossimo a 0 e più la proporzione dei fori nel materiale è elevata.

Qui viene sostituito il problema di ottimizzazione discreta originale di tipo 0 o 1 (ad ogni punto dello spazio, si ha un vuoto o del materiale), ad un nuovo problema di ottimizzazione *continua*, dove la variabile da ottimizzare, la densità di materiale, varia all'interno dell'intervallo [0,1].

L'idea di base è quella di trasformare un problema di ottimizzazione della forma in un problema di ottimizzazione della densità del materiale.

Con questo nuovo approccio, non si è più prigionieri della parametrizzazione delle forme proposte da Hadamard: la topologia potrà variare e i fori potranno apparire o scomparire a seconda della variazione di densità del materiale composito.

Osserviamo che la densità del materiale non è sufficiente a caratterizzare completamente un materiale composito: per una data densità, conta anche la forma dei fori per valutare le proprietà effettive del materiale. Ad esempio, la rigidità equivalente di un materiale composito non sarà la stessa per una struttura laminata rispetto ad una struttura a nido d'ape. Il metodo di omogeneizzazione propone, pertanto, di ottimizzare non solo la densità del materiale, ma anche la microstruttura (la forma dei fori) del materiale composito.

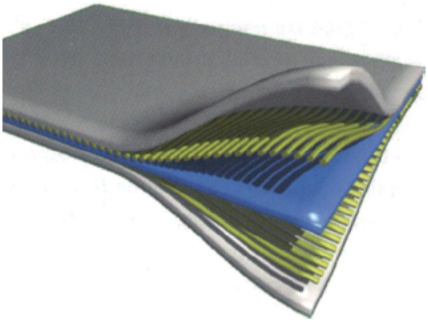
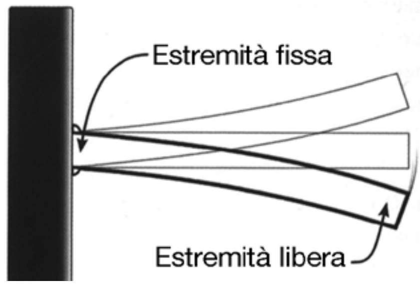


Figura 2. Esempio di materiale composito.



Un ulteriore vantaggio del metodo di omogeneizzazione è che la sua soluzione numerica è più veloce: essa richiede un minor numero di calcoli. Una ragione di questo è che, per cambiare la densità in un dato punto, l'algoritmo utilizza solo i valori di densità nel suo intorno, e non necessita dei valori dei punti più distanti, come invece avviene con il metodo di Hadamard.

La mensola ottimale

Ecco un problema classico di ottimizzazione della forma, quello della *mensola ottimale*. In questo caso, il bordo sinistro della mensola è fissato, mentre una forza verticale è applicata al centro del bordo di destra.

Disegniamo la densità di materiale, rappresentata da un livello di grigio (il nero corrisponde ai pieni, il bianco ai vuoti). La soluzione ottimale presenta grandi aree di grigio, corrispondente al materiale composito, che è difficile interpretare come una forma. Per trovare una forma "classica" che si approssimi alla forma *composita* ottimale, una soluzione è quella di "penalizzare" i materiali compositi nel processo di ottimizzazione numerica, vale a dire aggiungere dei vincoli all'algoritmo, affinché la soluzione ottimale sia costituita da zone o piene o vuote, piuttosto che da composite.

Il risultato è impressionante: la zona composita si trasforma in un reticolo di barre, che ricorda molto le strutture dell'ingegneria civile o meccanica.

Questi metodi di ottimizzazione della forma geometrica e topologica sono utilizzati

Figura 4. Mensola composita ottimale (il livello di grigio indica la densità del materiale).

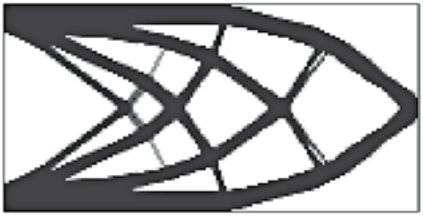


Figura 5. Mensola ottimale dopo la penalizzazione e senza compositi.

quotidianamente nei settori automobilistico e aerospaziale, per esempio quando si tratta di trovare la forma di una struttura che sia al contempo rigida e leggera. Molte parti meccaniche (bracci di sospensione, guide, ecc.) nelle automobili o negli aerei sono alleggeriti in questo modo attraverso l'ottimizzazione, al fine di ridurre il consumo di carburante.

Figura 6. Ottimizzazione di un ponte.

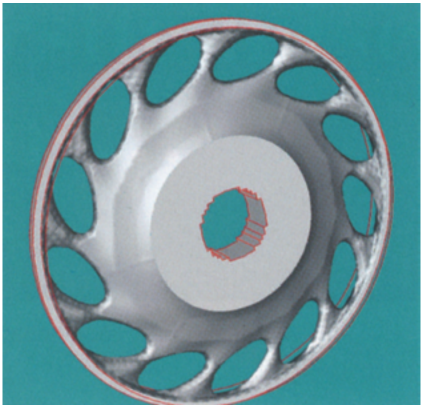


Figura 7. Ottimizzazione del pignone di una ruota.

Per saperne di più

Sito web:
http://www.cmap.polytechnique.fr/~optopo

Allaire G., (2007). *Conception optimale de structures*, Collection Mathématiques et Applications, vol. 58, Springer Verlag.

Henrot A., Pierre M., (2005). *Variation et optimisation de formes*, Collection Mathématiques et Applications, vol. 48, Springer Verlag.

Hildebrandt S., Tromba A., (1986). *Mathématiques et formes optimales*, *Pour la Science*, Belin, Paris.

La biodiversità tradotta in equazioni... o quasi

Sylvie Méléard, *professoressa presso École polytechnique*

Predire l'evoluzione di una popolazione animale su un lungo periodo, conoscere il funzionamento di un ecosistema, capire il vantaggio della riproduzione sessuale per la sopravvivenza delle specie ... i problemi che nascono dallo studio della biodiversità sono complessi e la loro risoluzione utilizza sofisticati strumenti matematici.

"After years, I have deeply regretted that I did not proceed far enough at least to understand something of the great leading principles of mathematics: for men thus endowed seem to have an extra-sense".

(Dopo anni, ho un profondo rimpianto per non aver progredito abbastanza nel capire almeno i grandi principi essenziali della matematica: gli uomini dotati di tale capacità di comprensione sembrano possedere un sesto senso.)

Charles Darwin,
Autobiografia

Può sembrare sorprendente associare la matematica, spesso vista come una successione di formule fisse, e la biodiversità, associata alla vita brulicante degli ecosistemi. Tuttavia, la complessità estrema, che si manifesta appena si osserva la biodiversità, giustifica l'interesse della matematica e dei matematici per questo campo. La sfida è quella di semplificare in modo intelligente i fenomeni al fine di poterli tradurre in equazioni e costruire così dei modelli matematici innovativi. L'obiettivo è quello di dedurre quantità calcolabili quali, ad esempio, il tempo di estinzione di una specie, la stima dell'abbondanza di una specie, la velocità d'invasione di un parassita, oppure di sviluppare algoritmi di simulazione, che sono strumenti indispensabili di sperimentazione

teorica per poter prevedere determinati comportamenti. Ciò è particolarmente vero per le grandi scale temporali dell'evoluzione, per le quali non esistono osservazioni dirette. Le simulazioni sono anche un supporto visivo che facilita lo scambio scientifico con i biologi. La modellizzazione infine può essere utilizzata per costruire degli strumenti statistici utili a prevedere e quantificare diversi scenari della biodiversità alla luce dei dati osservati.

Nel XIX secolo e all'inizio del XX secolo, molti modelli, soprattutto deterministici, vale a dire in cui la stessa causa provoca sempre lo stesso effetto, sono stati sviluppati per studiare la dinamica delle popolazioni (Malthus, Verhulst, Lotka, Volterra ...). Altri modelli più probabilistici, vale a dire, che coinvolgono il caso, sono stati proposti per studiare la genetica delle popolazioni (Fisher, Haldane, Wright ...). Negli ultimi decenni, i biologi si sono concentrati principalmente sulla biologia molecolare e sulla grande massa di dati che le nuove tecnologie hanno permesso loro di ottenere. Metodi di sequenziamento hanno fornito una massa formidabile di informazioni da filtrare, classificare, ordinare e analizzare. Con questi dati, i biologi sentono ancora di più la necessità di creare modelli che possano dare a queste rilevazioni una coerenza e consentire di ricavare previsioni sul comportamento biologico ed ambientale.

La complessità dei problemi relativi alla biodiversità richiede, quindi, sofisticati strumenti matematici. Allo stesso tempo, i matematici possono attingere da questi problemi motivazioni e ispirazioni per creare nuovi modelli e teoremi di grande ricchezza, che permetteranno di comprendere alcuni fenomeni ecologici.

Alcune questioni biologiche

Ecco alcuni esempi di domande di biologia alle quali i matematici cercano di dare delle risposte:

– Qual è il comportamento di una popolazione nel tempo? Si stabilizzerà o si estinguerà? Se si estinguerà, in quanto tempo?

– Le specie coabitano e interagiscono in una competizione per le risorse all'interno di una dinamica ospite-parassita o di una catena alimentare, o diversamente collaborano per sopravvivere. Come si evolvono questi sistemi?

– Qual è l'impatto della migrazione e della frammentazione degli habitat sulla biodiversità?

– Qual è il vantaggio della riproduzione sessuale per la biodiversità e la sopravvivenza delle specie? Anche se la riproduzione sessuale apporta maggiore variazione genetica attraverso la ricombinazione, la nutrizione dei maschi, l'accoppiamento, l'esposizione ai predatori durante il sesso, sono fattori che rendono più favorevole la riproduzione asessuata.

– Qual è l'impatto dei cambiamenti ambientali sulla biodiversità?

– Qual è il legame tra il DNA delle popolazioni, le loro interazioni e la loro evoluzione?

Il concetto di specie e modelli probabilistici

Per comprendere la biodiversità, dobbiamo prima capire che cos'è una specie.

Il concetto di specie si è evoluto nel tempo. Già Maupertuis (1698-1759), matematico e naturalista, osservò la comparsa di mutazioni casuali che si sviluppano e formano una nuova popolazione, "Vediamo apparire razze di cani, piccioni, canarini che non erano presenti in precedenza in natura. Inizialmente essi sono stati solo frutto del caso; il ripetersi delle caratteristiche nelle generazioni ha creato le specie." Nel lavoro di Lamarck (1744-1829), l'immagine lineare della grande catena degli esseri viventi è sostituita da un "albero con più rami." Fino a che, nella sua famosa opera *Sull'origine delle specie per mezzo della selezione naturale*, Darwin (1809-1882) ha introdotto l'idea della selezione naturale e dell'albero delle specie, con la perdita di specie per selezione del più adatto e l'apparizione di nuove, come le varie specie di fringuelli che egli osservò sulle isole Galapagos.

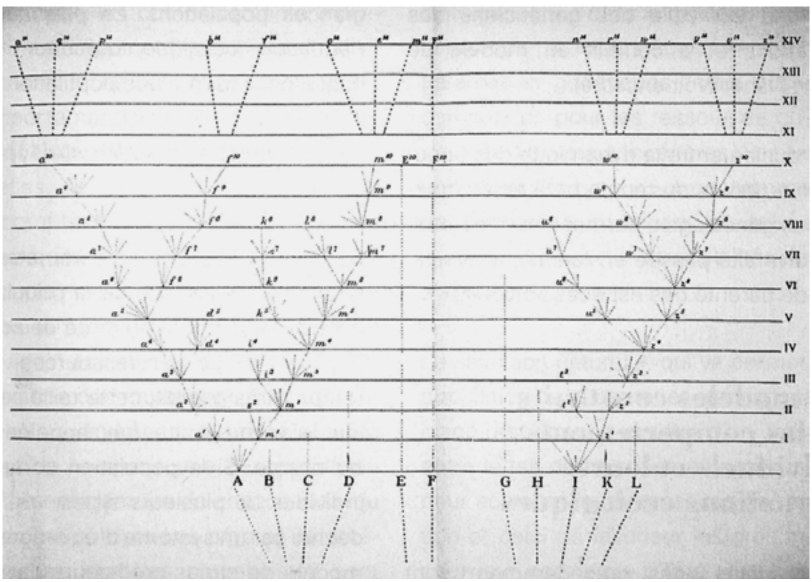


Figura 1. Albero delle specie con l'estinzione di certe specie e l'apparizione di nuove, secondo C. Darwin, L'origine delle specie per mezzo della selezione naturale.

È questa dinamica della biodiversità che vogliamo capire meglio, attraverso modelli matematici e, in particolare, modelli probabilistici che tengano conto della nascita e della morte in modo aleatorio degli individui e che si chiama generalmente *processo di ramificazione* o *processo di nascita e morte*. Il prototipo è il processo di Galton-Watson (vedi riquadro pag. 75), che può essere usato per calcolare la probabilità di estinzione di una specie.

Un approccio inverso, che la scoperta del DNA ha fortemente stimolato, è quello di ricostruire la storia di una popolazione. Dato un gruppo di individui, a che epoca risale il loro più recente antenato comune? Come risalire agli antenati di un individuo? Questo approccio è quello degli studiosi della genetica delle popolazioni e ha portato al modello di Wright-Fisher (vedi riquadro pag. 76).

Questa dualità tra la dinamica della popolazione nel tempo e la ricostruzione dell'albero genealogico, permette di stimare la biodiversità passata in funzione delle parentele tra le specie attuali.

I modelli basati sui comportamenti individuali e le loro interazioni biologiche

I modelli di cui sopra sono matematicamente molto semplici. Sebbene il processo Galton-Watson e i modelli di Wright-Fisher giochino un ruolo fondamentale – la loro semplicità permette di fare calcoli e di fornire alcune risposte quantitative – non tengono conto della diversità genetica e delle interazioni biologiche. Inoltre, le ricerche attuali dei biologi teorici e dei matematici hanno lo scopo di presentare la diversità genetica (DNA) e l'interazione (competizione, cooperazione, predazione) nei parametri demografici dei modelli.

La ricerca attuale si propone di tenere conto della diversità genetica e dell'interazione.

I primi modelli in cui si è introdotta la competizione sono stati dei modelli deterministici, i quali descrivono il comportamento di grandi popolazioni. Il più famoso è *l'equazione logistica*, che descrive lo sviluppo di una popolazione:

$$n'(t) = n(t)(b - cn(t))$$

In questa equazione, il parametro b descrive il tasso di crescita della popolazione e c misura la competizione tra due individui per le risorse condivise. Per tempi t molto grandi, questa grandezza si stabilizza sul valore limite b/c, chiamata *capacità di carico*. Se la popolazione comprende individui di tipo diverso, è necessario impiegare un sistema di equazioni. Un esempio è il modello di predatore-preda: se $n_1(t)$ e $n_2(t)$ sono il numero delle prede e dei predatori al tempo t allora abbiamo le equazioni:

$$n'_1(t) = an_1(t) - bn_1(t)n_2(t)$$

$$n'_2(t) = -cn_2(t) + dn_1(t)n_2(t)$$

In questo modello si osserva un comportamento ciclico di queste dinamiche.

Nel caso di un *continuum* di tipologie (grandezza alla nascita, l'età alla maturità ...), la dinamica di una popolazione di grandi dimensioni sarà descritta da equazioni più complicate (equazioni alle derivate parziali o integro-differenziali) che mettono in relazione le tipologie degli individui con il tempo.

Questi modelli deterministici sono modelli macroscopici (vale a dire che considerano il comportamento globale di un insieme di individui) e non tengono conto del comportamento individuale, né della nascita casuale di individui mutanti. Essi non sono rappresentativi del comportamento di piccole popolazioni nelle quali le fluttuazioni aleatorie sono essenziali. Inoltre non possono spiegare tutte le scale temporali alle quali dobbiamo lavorare: per esempio, il rapporto tra scala biologica (il tempo di nascita e di morte degli individui) e quello dell'evoluzione (l'apparizione e la fissazione delle mutazioni). Per integrare questi eventi nel modello, è importante ricorrere ad una descrizione del comportamento aleatorio di ogni individuo, tenendo conto delle sue peculiarità. Modelli probabilistici sono quindi delle generalizzazioni del processo Galton-Watson.

Esempio di una popolazione caratterizzata dal peso degli individui alla nascita

In questo esempio, gli individui sono caratterizzati dal peso x alla nascita, che varia ad esempio tra 0 e 4. Il tempo che un individuo di peso x impiega a riprodursi è una variabile aleatoria di valore atteso $b(x)$. Possiamo supporre ad esempio, che valga $b(x) = 1/(4-x)$ il che riflette il fatto che un individuo più grande necessita di più energia per nutrirsi sottraendola alla riproduzione.

Ad ogni riproduzione, l'individuo trasmette ereditariamente il suo tipo x, ma può succedere che il suo discendente muti in un individuo un tipo di y. Questo si riprodurrà a sua volta generando una sua sottopopolazione di tipo y. Le mutazioni introducono, così, la variabilità genetica nel modello. Inoltre, la morte dell'individuo dipende dai parametri genetici, ma anche dalla competizione per le risorse che deve sostenere con i suoi consimili. Ad esempio, un individuo di tipo $y > x$ consumerà più risorse, a scapito di quelle di x, che in questo modo vede diminuire le sue possibilità di sopravvivenza.

La selezione naturale, che permetterà alla popolazione di adattarsi e portare avanti la sua evoluzione, è il risultato di un compromesso tra il fatto di favorire i piccoli individui per accrescere la loro capacità riproduttiva e quella di favorire i grossi individui per consentire loro di essere più forti nella competizione per le risorse. È questo tipo di ottimizzazione evolutiva che i matematici cercano di capire, così come i meccanismi matematici che possono spiegare la comparsa di nuove specie.

Il modello sopra descritto può spiegare la diversità dei fringuelli di Darwin attraverso la dimensione del loro becco. La figura mostra la simulazione di un modello (vedi figura 2), che spiega la comparsa di quattro specie da una singola specie su di una scala di tempi molto lunghi.

Naturalmente è totalmente illusorio poter ridurre la biodiversità a delle semplici equazioni, ma i modelli matematici possono fornire una nuova prospettiva e un punto di vista svincolato e obiettivo degli ecosistemi.

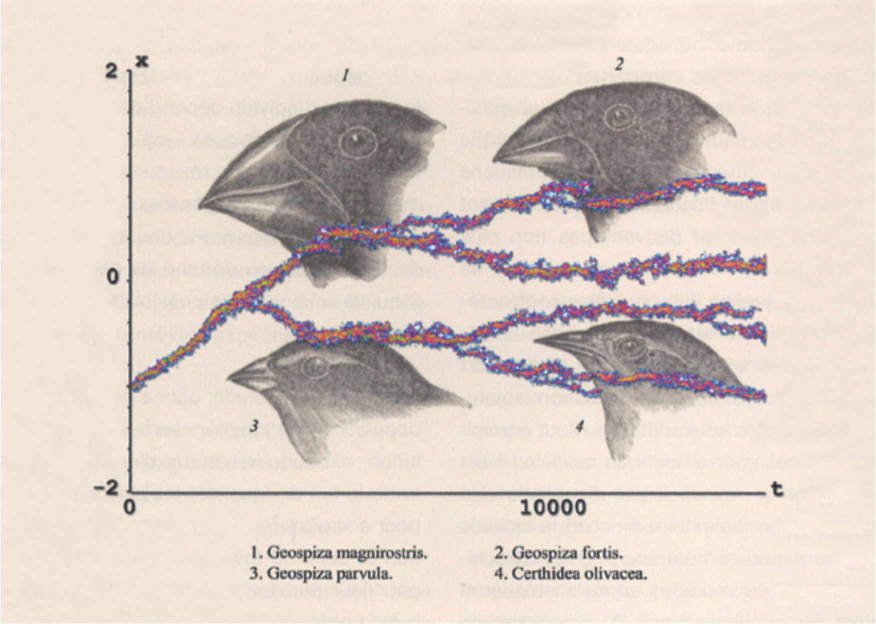


Figura 2. Simulazione del modello dei fringuelli di Darwin.

Il processo di Galton-Watson

Il processo di Galton-Watson è una sequenza di variabili aleatorie (Z_n) che descrivono l'evoluzione di una popolazione che, ad ogni generazione si riproduce in modo casuale. Se Z_n è la grandezza della popolazione alla n-esima generazione, allora $Z_{n+1} = X_1 + \dots + X_{Z_n}$, dove X_i è il numero di figli del'i-esimo individuo. Le variabili X_i seguono tutte la stessa legge X, in particolare il numero medio m di figli per individuo è costante. La probabilità di estinzione si ottiene, allora, come limite di una successione ricorsiva (nella quale il termine generale dà la probabilità di estinzione alla generazione n) ed è una soluzione dell'equazione $g(p) = p$, dove $g(s)$ è la media de s^X. Si può dimostrare che se m è un numero maggiore di 1, la popolazione può esplodere esponenzialmente, mentre se m è inferiore o uguale ad 1, la popolazione si estingue. Ad esempio, sapendo che il numero medio dei piccoli delle balene nere è stimato a 0,976 e l'abbondanza delle femmine è stimata a 150 nel 1994, si può dedurre usando il modello che ci sia il 99% di possibilità che queste balene siano completamente scomparse nel 2389.

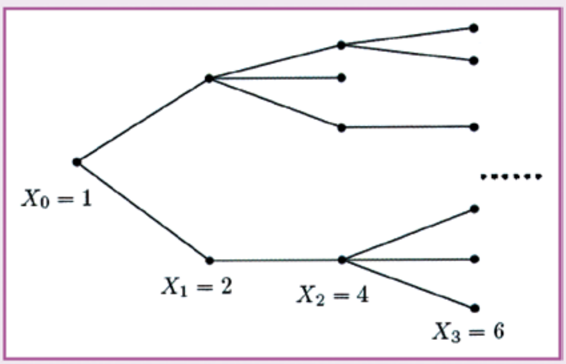


Figura 3. Una realizzazione del processo di Galton-Watson.

Il modello di Wright-Fisher

In questo modello, la grandezza della popolazione N è costante e, ad ogni generazione, ogni individuo sceglie il suo antenato con probabilità uniforme tra gli individui della generazione precedente. In questo modo, la probabilità che due individui abbiano lo stesso genitore è $1/N$, che tende a 0 quando N tende all'infinito. È necessario, quindi, modificare la scala dei tempi per osservare la genealogia di una grande popolazione e considerare i tempi di generazione proporzionali ad N. In questo caso, si costruisce un oggetto limite chiamato *coalescente di Kingman*, che è un oggetto geometrico probabilistico che descrive le linee di discendenza.

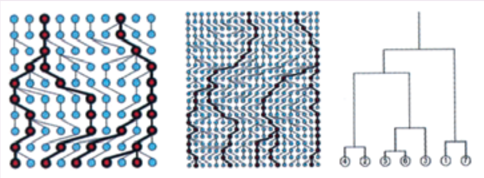


Figura 4. Il coalescente Kingman come limite di scala degli alberi genealogici (B. Mallein, Culture Math, ENS Ulm, 2011)

Per saperne di più

Bacaer N., (2008). *Histoires de mathématiques et de populations.* Le sel et le fer, Cassini.

Haccou P., Jagers P., Vatutin VA, (2005). *Branching Processes Variation, growth: and extinction of populations*, Cambridge University Press.

Istas J., (2000). *Introduction aux modélisations mathématiques pour les sciences du vivant*. Springer.

Mallein B., (2011). *Généalogie de populations: le coalescent de Kingman*, Culture Math, ENS Ulm.

http://www.math.ens.fr/culturemath/articles-ens/mallein11/coalescent-de-kingman.html

Tangente Hors série n° 42 (2011). *Mathématiques et biologie – L'organisation du vivant*.

Darwin online http://darwin-online.org.uk/content/frameset?pageseq=60&itemID=F1497&viewtype=side

Il restauro dei vecchi film

Julie Delon, *ricercatrice CNRS presso Telecom ParisTech*
Agnès Desolneux, *direttrice di ricerca CNRS presso Ecole Normale Supérieure de Cachan*

Lo sfarfallio fa parte dei difetti che comunemente colpiscono le pellicole danneggiate. Attraverso alcuni casi particolari, vedremo come la matematica aiuta a creare degli algoritmi per correggere automaticamente le imperfezioni dei vecchi film.

L'emergere delle tecnologie digitali offre oggi l'accesso ad una parte importante del patrimonio cinematografico. Tuttavia, il processo di digitalizzazione deve essere accompagnato dal restauro dei molti difetti che alterano i film e i video. Tra questi inconvenienti citiamo in particolare lo *sfarfallio* (chiamato anche *flicker* per usare il termine inglese), che si traduce in variazioni innaturali dell'intensità dell'immagine, graffi, viraggio del colore, *macchie* (piccoli punti dovuti a granelli di polvere o alla perdita di pezzi di gelatina sulla pellicola), ecc.

La varietà dei difetti riscontrabili nei film rende particolarmente difficile il loro restauro. Un film di 90 minuti contiene circa 130.000 immagini. Per restaurarlo completamente in un tempo ragionevole e per evitare un eccessivo investimento umano, diventano indispensabili degli algoritmi di restauro rapidi e il più automatizzati possibile. Lo sviluppo di tali algoritmi è ancora più determinante con l'attuale tendenza verso l'uso di immagini a risoluzione sempre più alta (ad esempio le TV HD o il supporto Blu-ray) che accentuano ogni minimo difetto nelle immagini.

Lo sfarfallio

Concentriamoci su di un difetto molto comune nei vecchi film: lo sfarfallio (o *flicker*). Questo difetto consiste in variazioni innaturali di con-

Figura 1. Tre immagini dal film "Charlot fa una cura" (1917) di Charlie Chaplin. L'immagine centrale è molto più scura rispetto rispetto alle altre due. È questa mancanza di luminosità, chiamata sfarfallio, o flicker, che dà l'impressione che il film "lampeggi" e che si cerca di correggere in modo automatico.

trasto tra un'immagine e l'altra del film: le immagini diventano artificialmente scure o chiare, alternativamente. Queste variazioni di contrasto possono essere dovute sia alla degradazione chimica della pellicola (che produce quindi delle zone più scure o più chiare durante la visualizzazione), ma anche ad un cambiamento nel tempo di esposizione (il tempo nel quale la pellicola è stata esposta alla luce) di un'immagine rispetto all'altra. Questo è particolarmente vero per i film realizzati all'epoca in cui i film venivano girati manualmente.

A differenza di altri difetti comunemente osservati nei film (graffi, polvere, ecc.), lo sfarfallio non fa apparire delle nuove strutture nelle immagini. La sua particolarità è, quindi, di essere trasparente, cioè non rilevabile su una singola immagine. Solo la visione dei fotogrammi successivi del film permette di rendersi conto della sua presenza. Pertanto, il ripristino di tale difetto non può essere fatto su ogni immagine in modo indipendente, ma è necessario utilizzare diverse immagini successive del film e cercare di "mediare il contrasto."

Per correggere questo difetto, i film sono prima digitalizzati, il che significa che la pellicola viene sottoposto a scansione, fotogramma per fotogramma, e che questo insieme di immagini digitali viene memorizzato su di un computer. In generale, un secondo di film contiene 24 immagini. Un film di un'ora, una volta digitalizzato, quindi contiene 86 400 immagini digitali. Un'immagine digitale in bianco e nero è matematicamente modellata come una funzione definita su una griglia rettangolare di quadrati (chiamati pixel, contrazione di *picture element*) a valori nell'insieme dei numeri positivi. Il valore dell'immagine in un pixel è chiamato livello di grigio di tale pixel.

Cambiamento di contrasto

Una volta digitalizzata la pellicola, possiamo applicare a tutte le immagini quello che si chiama il *cambiamento di contrasto* (vedremo più avanti come operare per eliminare lo sfarfallio). Dire che un'immagine I subisce un cambiamento di contrasto significa che viene trasfor-

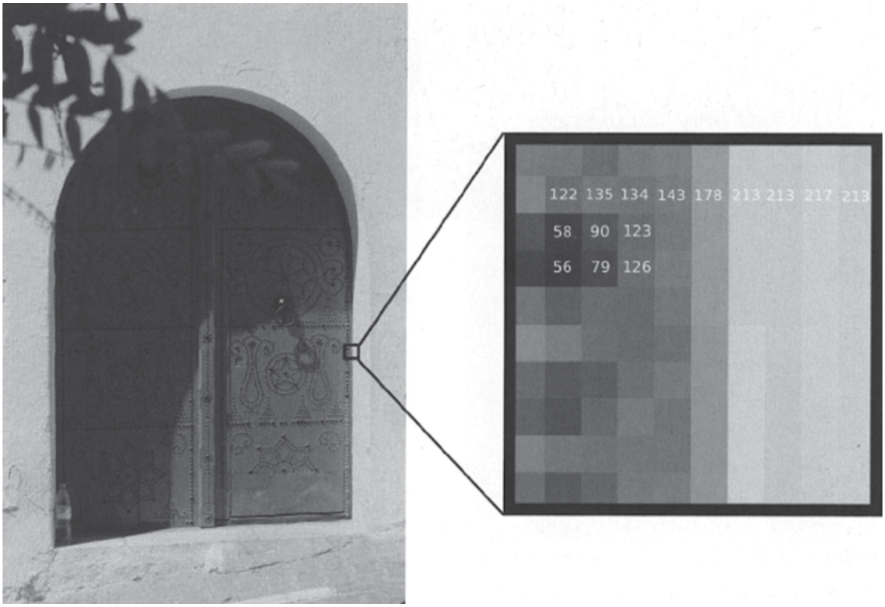


Figura 2. Una immagine digitale e un ingrandimento di una piccola parte di questa immagine, che mostra alcuni pixel e i loro livelli di grigio.

mata in $f(I)$ dove f è una funzione monotona crescente sull'insieme dei numeri positivi: ad ogni pixel di coordinate (x, y) viene cambiato il suo livello di grigio $I(x, y)$ nel valore $f(I(x, y))$ Così, due pixel aventi lo stesso livello di grigio nell'immagine I avranno ancora livelli di grigio identici nella nuova immagine $f(I)$ L'interesse e la necessità di utilizzare una funzione f crescente risiede nel fatto che conserva l'ordine dei livelli di grigio: se un pixel è più scuro di un altro pixel nell'immagine I, questa proprietà rimane vera nell'immagine $f(I)$. Di conseguenza, un cambiamento di contrasto non modifica il contenuto geometrico di un'immagine, vale a dire che si vede la stessa cosa nell'immagine prima e dopo un cambiamento di contrasto. Nell'immagine non appaiono nuovi oggetti.

Un cambiamento di contrasto non altera il contenuto geometrico di un'immagine, vale a dire che si vede la stessa cosa nella foto prima e dopo un cambio di contrasto.

I cambiamenti di contrasto usati per correggere lo sfarfallio nei film sono costruiti come segue. Supponiamo di voler ripristinare la n-esima immagine di un filmato, che indicheremo con I_n. Per ogni pixel (x, y) di I_n, guardiamo quello che è il rango R del pixel nell'immagine I_n,

Figura 3. Cambiamento di contrasto. L'immagine di destra è ottenuta applicando una variazione contrasto all'immagine a sinistra. Si può osservare che il contenuto geometrico dell'immagine non è cambiato.

in modo che tutti i pixel siano ordinati in maniera crescente in base al loro livello di grigio (si assume che questo rango sia uguale per pixel che hanno lo stesso livello di grigio). Adesso confrontiamo, nelle 10 immagini che precedono I_n e nelle 10 che lo seguono all'interno del filmato (vedi riquadro *Il numero di immagini*), quali sono i livelli di grigio dei pixel aventi lo stesso rango R (o il rango più vicino ad R) in queste immagini. Questi pixel hanno valori di grigio che possono essere molto diversi da quello di $I_n(x, y)$. Pertanto prendiamo la media di tutti questi livelli di grigio (21 valori in totale), il che determina il nuovo valore $f(I_n(x,y))$ del pixel (x, y) per l'n-esima immagine del filmato. Tale correzione si basa sulla seguente osservazione: se per un oggetto del filmato il suo livello di grigio può variare molto, il rango dei pixel che contiene invece cambierà di poco da un'immagine all'altra. Ad esempio, i pixel del livello della mediana (quelli per cui il numero di pixel dell'immagine più scuri uguaglia quello dei pixel più chiari) è probabile che corrispondano agli stessi oggetti in immagini diverse del film (vedi riquadro *L'istogramma cumulativo dei livelli di grigio*).

Abbiamo così ottenuto una *equalizzazione di contrasto* tra le immagini del film, che elimina l'effetto dello sfarfallio.

Figura 4. Le tre immagini del film Charlie Chaplin "Charlot fa una cura" dopo il restauro automatico del contrasto con il metodo descritto in questo articolo. Queste tre immagini hanno adesso una luminosità simile, e l'effetto sfarfallio (che produceva un lampeggiamento) del film è scomparso.

Riferimenti

Questo articolo è ispirato a quello che abbiamo scritto per il sito Image des Mathématiques:

Delon J., Desolneux A., (2011) *Papillonnage et mathématiques des images – Images des Mathématiques, CNRS*.

En ligne, URL: http://images.math.cnrs.fr/Papillonnage-et-mathematiques-des.html

Handbook of Mathematical Models in Computer Vision, Edited by Nikos Paragios, Chen Yunmei, et Olivier Faugeras, Springer, 2005

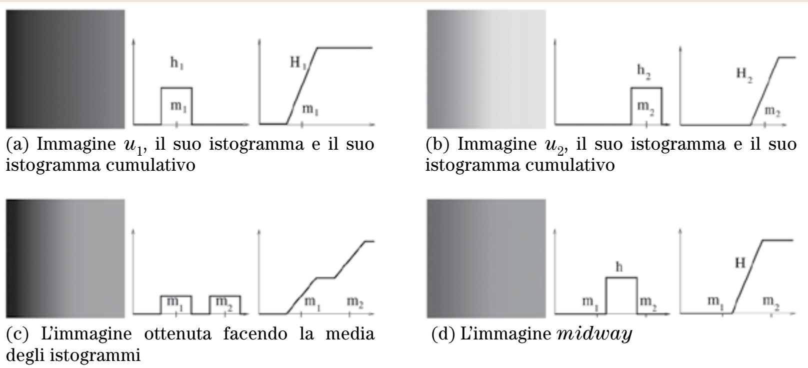
Handbook of Mathematical Methods in Imaging, Edited by Otmar Scherzer, Springer, 2011

Il numero di immagini

Il numero di immagini utilizzate per il restauro (abbiamo preso qui le 10 immagini precedenti e le 10 immagini seguenti) è una scelta euristica che può essere modificata. Più questo numero è grande, più si tiene conto di un gran numero di immagini nel film per ripristinare l'immagine corrente, e più si eliminano le fluttuazioni dovute allo sfarfallio. D'altra parte, prendere troppe immagini non ha necessariamente senso se il film ha molti movimenti di camera o oggetti che si spostano.

L’istogramma cumulativo dei livelli di grigio

Il cambiamento di contrasto applicato all’immagine I_n del filmato si può interpretare in modo accurato mediante il concetto di istogramma cumulativo (chiamato anche funzione di distribuzione) dei livelli di grigio dell’immagine. L’istogramma cumulativo dei livelli di grigio di un’immagine è definito come la funzione che a ciascun valore del livello di grigio (cioè, generalmente ad ogni numero compreso tra 0 e 255) associa il numero di pixel nell’immagine avente un livello di grigio inferiore a questo valore. È quindi una funzione crescente del livello di grigio. L’operazione che consiste nel fare la media dei livelli di grigio dei pixel dello stesso rango R nelle 10 immagini che precedono e nelle 10 che seguono cronologicamente I_n nel filmato, equivale infatti a cambiare l’istogramma cumulativo di I_n e a fare la media armonica (il reciproco della media aritmetica dei reciproci) degli istogrammi cumulativi delle 21 immagini considerate. Altri tipi di medie possono essere prese in considerazione, in particolare la media aritmetica, ma si avrebbe allora l’apparizione di artefatti, come illustrato nella figura 5. In effetti, si può dimostrare che la media armonica è l’unica che garantisce, per un filmato costituito soltanto da un’immagine fissa, di effettuare correttamente la media del contrasto.



(a) Immagine u_1, il suo istogramma e il suo istogramma cumulativo

(b) Immagine u_2, il suo istogramma e il suo istogramma cumulativo

(c) L’immagine ottenuta facendo la media degli istogrammi

(d) L’immagine *midway*

Figura 5. (a) Immagine di un gradiente scuro, con il suo istogramma e il suo istogramma cumulativo. (b) Immagine di un gradiente chiaro, con il suo istogramma e il suo istogramma cumulativo. (c) L’immagine ottenuta se avessimo usato la media aritmetica per la correzione del contrasto: si ottiene una discontinuità nel gradiente. (d) L’immagine ottenuta con la media armonica degli istogrammi cumulativi: il risultato è in linea con quello che ci si aspetta - si ottiene un gradiente di livello “medio” di grigi.

Criptaggio e decriptaggio: comunicare in tutta sicurezza

Jean-Louis Nicolas, *professore presso Université Claude Bernard Lyon 1*

Christophe Delaunay, *professore presso Université Franche-Comté*

La protezione delle nostre carte bancarie, così come altre procedure crittografiche comunemente utilizzate, si basa sull'impossibilità pratica di fattorizzare numeri molto grandi. La supremazia di questi metodi, la cui affidabilità è costantemente messa in discussione dai progressi informatici, potrebbe in futuro essere scalzata da altre tecnologie.

Nel marzo 2000 un grosso titolo era sulla prima pagina di tutti i giornali: "Attenzione alla sicurezza delle carte bancarie". Cosa era successo? In Francia, il codice segreto delle carte magnetiche era protetto dal 1985 grazie ad un metodo di criptaggio che usava un grande numero N costituito da 97 cifre. Questo numero N doveva essere il prodotto di due grandi numeri primi, vale a dire di numeri che, come ad esempio il 7 o il 19, sono divisibili solo per 1 e per se stessi. Il codice segreto di una carta bancaria è costituito esattamente da questa coppia di numeri primi; il loro calcolo a partire da N era praticamente impossibile negli anni '80. Ma con l'aumento della potenza dei calcolatori e il miglioramento dei metodi matematici, la grandezza dei numeri N, di cui non si possono calcolare i fattori primi in tempi ragionevoli, ha oltrepassato il centinaio di cifre negli anni '90 (il record attuale, ottenuto nel dicembre 2009, è la fattorizzazione di un numero di 232 cifre). Così, un astuto informatico, Serge Humpich, aveva potuto trovare i due numeri primi ultra segreti il cui prodotto valeva N e li aveva utilizzati per fabbricare carte bancarie false. Per garantire la sicurezza dei nostri piccoli rettangoli di plastica, l'organismo di gestione delle carte bancarie è stato costretto a costruire immediatamente dei nuovi numeri N nettamente più grandi.

Il sito web di Paul Zimmerman (vedi la bibliografia a fine articolo) fornisce un aggiornamento dei vari record di fattorizzazione di numeri interi.

La dimensione del numero N i cui fattori primi possono essere calcolati in un tempo ragionevole ha superato il centinaio di cifre negli anni '90.

La crittografia moderna, un incrocio fra matematica e informatica

Questa vicenda illustra l'importanza considerevole che riveste oggi la scienza del criptaggio, vale a dire della codifica dei messaggi

con lo scopo di renderli illeggibili ad occhi indiscreti.

Criptare e decriptare dei messaggi segreti è un'attività vecchia di molti secoli, addirittura millenni. E questa attività ha largamente superato i confini degli ambienti strettamente diplomatici o militari per investire interi settori dell'universo delle comunicazioni civili: procedure d'autenticazione, transazioni bancarie, commercio elettronico, protezione di siti e archivi informatici, ecc.

La crittografia ha conosciuto vari sviluppi nel corso degli ultimi decenni. Di conseguenza, essa è divenuta una scienza complessa, dove i progressi sono generalmente opera di specialisti con una formazione matematica e informatica.

Questa specializzazione si è manifestata a partire dalla seconda guerra mondiale. Oggi sappiamo che la decriptazione da parte degli alleati dei messaggi codificati dalla famosa macchina tedesca "Enigma" ha giocato un ruolo importante in questo conflitto. Un eminente matematico britannico, Alan Turing, uno dei padri dell'informatica teorica, ha apportato un contributo essenziale a questa decriptazione.

Negli anni '70 la crittografia ha conosciuto una piccola rivoluzione: l'invenzione della crittografia a "chiave pubblica" con il metodo RSA. Di che cosa si tratta? Fino ad allora i corrispondenti che volevano scambiarsi messaggi segreti condividevano una chiave segreta e il rischio di intercettazione di questa chiave da parte del nemico era notevole. Il protocollo RSA, chiamato così in onore dei suoi inventori

(Ronald Rivest, Ad Shamia e Leonard Adleman) ha risolto questo problema. Questo metodo usa due chiavi: una chiave di criptazione pubblica - che può essere conosciuta da tutti - e una chiave di decriptazione, che rimane segreta. Essa si basa sul principio (utilizzato in seguito per proteggere le carte bancarie, come abbiamo visto in precedenza) che è possibile costruire dei grandi numeri primi (di cento, mille cifre, e oltre), ma che è estremamente difficile ritrovare i fattori primi p e q di un grande numero N, che sia $N = p \times q$, conoscendo solo N. Schematicamente, la conoscenza di N rimanda a quella della chiave pubblica di criptazione, mentre la conoscenza di p e q rimanda a quella della chiave segreta di decriptazione.

Evidentemente, se qualcuno trovasse un metodo per decomporre rapidamente nei loro fattori primi dei grandi numeri, il protocollo RSA diventerebbe inefficace. Ma potrebbe anche succedere che i matematici provino che tale metodo non esiste, cosa che rafforzerebbe la sicurezza del protocollo RSA. Questi sono degli argomenti di ricerca fondamentali.

I metodi che, come il protocollo RSA, fanno intervenire elaborate teorie dei numeri, ci danno una grande lezione: le ricerche matematiche (sui numeri primi specialmente) fatte disinteressatamente, possono rivelarsi, anni o decenni più tardi, cruciali per svariate applicazioni; e questo in maniera imprevedibile.

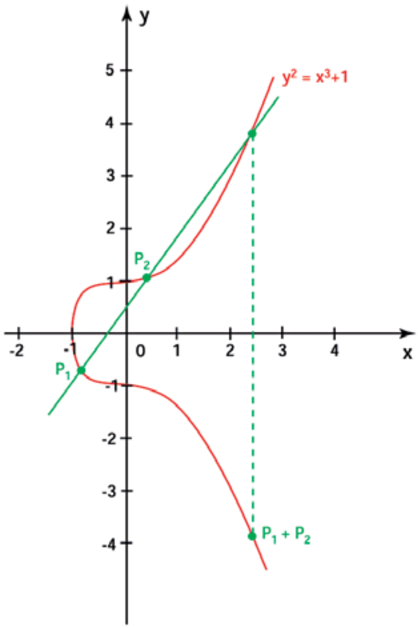
Nel suo libro "*Apologia di un matematico*", il grande teorico dei numeri britannico G. H. Hardy (1877-1947), il quale era un fervente pacifista, si vantava di lavorare in un campo perfettamente puro, l'aritmetica, e di non aver mai fatto niente che potesse essere considerato "utile". I suoi lavori erano forse "inutili" nella sua epoca. Oggi questo non è più vero.

Curve ellittiche: la geometria algebrica al servizio degli agenti segreti

Tutto ciò non riguarda soltanto la teoria dei numeri. Altri campi della matematica pura, considerati privi di applicazione, contribuiscono alla sicurezza del criptaggio. Metodi crittografici promettenti basati su principi vicini a quelli del protocollo RSA sono apparsi nel corso degli ultimi anni. È così per il metodo detto del *logaritmo discreto*.

Questo è servito a sua volta a concepire dei metodi che si basano sulle proprietà delle *curve ellittiche*. Non si tratta di curve aventi la forma di una ellisse, ma di curve il cui studio è iniziato nei primi anni del secolo XIX per risolvere il difficile problema del calcolo del perimetro di una ellisse. Queste curve, per le quali le

coordinate (x, y) dei loro punti verificano una equazione del tipo $y^2 = x^3 + ax + b$, hanno interessanti proprietà – il cui studio fa parte della *geometria algebrica* – un settore molto vasto della matematica attuale. Per esempio, con l'aiuto di una appropriata costruzione geometrica, è possibile definire una somma fra i punti di una curva ellittica (vedi figura sottostante).



Il grafico della curva ellittica di equazione $y^2 = x^3 + 1$. Le curve ellittiche hanno una notevole proprietà: si possono "addizionare" i loro punti secondo il procedimento indicato sul disegno. La "somma" così definita rispetta le consuete leggi aritmetiche; ad es. $(P_1 + P_2) + P_3 = P_1 + (P_2 + P_3)$. Certi metodi attuali di crittografia necessitano delle curve ellittiche e delle loro proprietà algebriche.

In generale, le curve ellittiche sono oggetti geometrici che possiedono delle proprietà aritmetiche – che sono sotto continua indagine – in grado di rendere un utile servizio alla crittografia. In tal modo è stato sviluppato un metodo crittografico detto *logaritmo discreto sulle curve ellittiche*. Curiosamente, le curve ellittiche forniscono anche un nuovo metodo per la fattorizzazione di numeri interi (tuttavia, dei metodi molto più avanzati sono necessari per ottenere i record attuali).

Le curve ellittiche possiedono delle proprietà aritmetiche che si possono impiegare in crittografia.

Il computer quantistico: lo strumento del futuro?

Un'altra direzione, totalmente diversa, è emersa di recente. Si tratta della crittografia quantistica. Che cosa significa? Da qualche anno, alcuni fisici e matematici hanno immaginato che un giorno sarà possibile realizzare un computer quantistico, vale a dire un calcolatore il cui funzionamento sfrutterebbe le bizzarre leggi della fisica quantistica, quelle che regnano nel mondo dell'infinitamente piccolo. Ci si è resi conto che un tale computer, se fosse realizzabile, sarebbe in grado di fattorizzare molto più rapidamente grandi numeri e renderebbe, quindi, del tutto inefficace il metodo RSA (in questo contesto, al congresso internazionale dei matematici a Berlino del 1998, Peter Shor, dei laboratori AT & T ha ottenuto il Premio Nevanlinna per i suoi lavori sulla fattorizzazione con i computer quantistici). Peraltro, alcune ricerche riguardo l'effettiva realizzazione pratica di un computer quantistico sono state pubblicate sulla rivista britannica *Nature* (vedi bibliografia). D'altro canto, alcuni ricercatori hanno elaborato alcuni protocolli di crittografia quantistica, vale a dire metodi di criptaggio che utilizzano degli oggetti (fotoni, atomi, ...) che obbediscono alle leggi quantistiche. Questi protocolli quantistici garantiscono (almeno in linea teorica) una sicurezza infallibile. Tutto ciò è in fase di elaborazione e potrebbe diventare operativo nel prossimo futuro (ad esempio, un cavo di comunicazione quantistica che collega le città di Ginevra e Losanna opera già da diversi anni ...).

Per saperne di più

Il sito di Paul Zimmerman sui records di fattorizzazione di numeri interi: www.loria.fr/~zimmerma/records/factor.html

Kahn D., (1980). *La guerre des codes secrets* (Interéditions).

Stern J., (Jacob O. 1998). *La science du secret*.

Singh S., (Lattès J.-C., 1999) *Histoire des codes secrets*.

Delahaye J.-P., (2000). *Merveilleux nombres premiers* (Belin/Pour la Science).

Stinson D. (Vuibert, 2001). *Cryptographie, théorie et pratique*.

Vandersypen L. M. K., *et al.*, (2001) *Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance*, Nature, vol. 414, p. 883-887.

Come e perché nuotare nel miele?

François Alouges, Guilhem Blanchard Sylvain Calisti, Simon Calvet, Paul Fourment Christian Glusa, Romain e Mario Leblanc Quillas-Saavedra,
rispettivamente professore e studenti presso École polytechnique

Il nuoto nei mezzi altamente viscosi come il miele è oggetto di ricerche in corso, che coinvolgono discipline diverse come la meccanica dei fluidi, la matematica applicata e la biologia. Ma perché interessarsi al nuoto nel miele? E quali sono le differenze tra nuotare nel miele e nell'acqua?

Chi vorrebbe fare il bagno nel miele? Nessuno. Chi si interessa al nuoto nel miele o, in generale, nei mezzi altamente viscosi, lo fa per studiare alcuni problemi, anche se in apparenza molto diversi tra loro, per i quali entrano in gioco gli stessi meccanismi. Infatti, il nuoto in mezzi altamente viscosi su scala umana (la scala macroscopica, dell'ordine di un metro) ha le stesse caratteristiche del nuoto in acqua su una scala molto piccola (scala microscopica, delle dimensioni nell'ordine del centesimo di millimetro). Ed è più facile realizzare esperimenti a grandezze normali nel miele piuttosto che esperimenti in miniatura nell'acqua!

Con questo artificio, possiamo studiare il nuoto in mezzi acquosi (vale a dire, a base di acqua) in scale molto piccole. E questa volta le applicazioni pratiche sono numerose. Ad esempio, i batteri sono dei microrganismi che si muovono e nuotano in liquidi simili all'acqua come il sangue. Comprendere meglio le loro capacità di movimento potrebbe permettere sia di impedire loro di muoversi, nell'ambito della lotta contro alcune malattie, sia, al contrario, di aiutarli a svilupparsi (ad esempio nel caso di batteri che costituiscono la flora intestinale che ci aiuta a digerire il cibo che mangiamo). Un altro esempio di applicazioni mediche ancora più evidente è quello dei micro-robot nuotatori. In effetti, se si potessero costruire dei minuscoli robot in grado di muoversi in ambienti come l'acqua o il sangue, potremmo usarli per il trasporto di farmaci fino alle cellule malate, effettuare interventi chirurgici sul corpo umano senza bisogno di usare il bisturi del chirurgo, oppure effettuare riparazioni di dimensioni microscopiche, inaccessibili con gli strumenti, seppure molto sofisticati, maneggiati dagli esseri umani.

Inerzia e viscosità

Quali sono le differenze tra il nuoto di Michael Phelps, 1m93, e quella del batterio Escherichia Coli (vedi figura 1), che misura pochi micrometri (millesimi di millimetro)? Perché non si può studiare il movimento di questi due esseri viventi nello stesso modo?

In generale, le due tipologie di effetti che si verificano quando ci si muove in un fluido sono quelli dell'inerzia e della viscosità. L'effetto dell'inerzia è quello che, per esempio, sperimentiamo all'interno di un aereo al momento del decollo quando ci sentiamo schiacciati contro il sedile mentre l'aereo accelera bruscamente. Questi effetti sono legati al fatto che le forze fisiche (attrito, peso, ecc) non agiscono direttamente sulla nostra velocità, ma piuttosto sulla nostra accelerazione.

È più facile realizzare esperimenti a grandezze normali nel miele piuttosto che esperimenti in miniatura nell'acqua!

La viscosità, invece, è provocata dalle interazioni tra le diverse molecole che compongono il fluido. Possiamo associarla alla resistenza che oppone un fluido quando si cerca di deformarlo.

Figura 1. Un campione di nuoto a confronto con E. Coli. Una differenza di.... grandezza.

Questi due effetti sono di intensità molto diverse a seconda della scala nella quale si opera. Per misurare la loro importanza relativa, si usa una quantità adimensionale chiamata numero di Reynolds. Il numero di Reynolds di un flusso è dato dalla formula $Re = \rho UL / \nu$, in cui ρ e ν rappresentano la densità e la viscosità del fluido, e L e U sono rispettivamente una lunghezza e una velocità di scorrimento caratteristiche (per esempio quelli del nuotatore): un numero di Reynolds molto grande (rispetto ad 1) significa che gli effetti inerziali sono molto più influenti di quelli viscosi, che quindi potremo trascurare. Se, al contrario, il numero Reynolds è molto piccolo, ci si limiterà a considerare la viscosità, trascurando l'inerzia. Ecco dunque qual è la differenza tra Michael Phelps e E. Coli: il numero di Reynolds che caratterizza il nuoto del campione olimpico è di circa 10 milioni, mentre quello che caratterizzano il nuoto di un batterio è dell'ordine di 0.00001.

Due flussi aventi un numero di Reynolds simile hanno caratteristiche simili, anche se corrispondono a situazioni molto diverse.

Due flussi aventi un numero di Reynolds simile hanno caratteristiche simili, anche se corrispondono a situazioni molto diverse, come ad esempio per il movimento di microrganismi nell'acqua, di oggetti della dimensione di un centimetro in un fluido viscoso (miele, silicone, vernice, ecc.) oppure il flusso di scorrimento di un ghiacciaio. In tutti e tre i casi, si ha un numero di Reynolds piccolo.

Il teorema della conchiglia Saint Jacques

Il nuoto con un basso numero di Reynolds, nel quale cioè domina l'effetto della viscosità, ha proprietà interessanti e talvolta inaspettate. Il moto del fluido è descritto da *equazioni di Stokes* che sono infatti una versione semplificata delle *equazioni di Navier-Stokes* rispetto alle quali l'inerzia viene trascurata. Hanno una caratteristica estremamente importante dal punto di vista matematico: sono lineari, vale a dire, la variazione di un parametro (pressione, velocità, etc.) provoca una variazione proporzionale di tutti gli altri parametri.

Questa linearità ha importanti implicazioni per il nuoto con basso numero di Reynolds, come la reversibilità: se un oggetto viene deformato in qualche modo e successivamente riportato alla sua forma originale seguendo precisamente il movimento contrario – nel qual caso si dice che il movimento è *inverso* – ritornerà allora precisamente nella sua posizione iniziale, cioè non subirà nessuno spostamento. Questo è il "teorema della conchiglia Saint Jacques", stabilito da E.M. Purcell nel 1977: queste conchiglie che nuotano aprendo e chiudendo il loro guscio non possono spostarsi in ambienti dove domina la viscosità.

Figura 2. Le conchiglie Saint Jacques non potrebbero spostarsi in un ambiente dominato dalla viscosità.

Per superare questa difficoltà, nel 1977 Purcell immaginò un sistema con due pinne in grado di nuotare con basso numero di Reynolds, azionandole in sequenza come mostrato schematicamente in figura 3. Si noti che nel movimento indicato non è presente un movimento inverso.

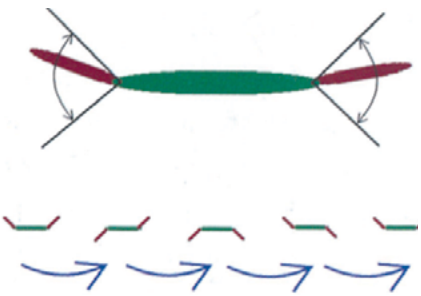


Figura 3. Il "Three-Link-Swimmer" di Purcell è un robot nuotatore con due pinne. Azionando in sequenza le due pinne come nel diagramma, il nuotatore progredisce lateralmente in un flusso dominato dalla viscosità.

Uno stile di nuoto ottimale

Il problema dello stile del nuoto è quello di cercare dei cambiamenti periodici di forma, delle bracciate, che producano, attraverso l'interazione con il fluido, uno spostamento del nuotatore. Il nuotatore (o il robot-nuotatore) può quindi concatenare le bracciate in modo da muoversi nel fluido. Matematicamente, questo problema equivale a cercare un *percorso chiuso* nell'insieme di tutte le forme possibili dell'oggetto che porti ad un cambiamento della sua posizione.

> *Il problema consiste nel trovare delle forme di nuoto che minimizzino l'energia meccanica spesa dal nuotatore per produrre un dato spostamento. Si tratta quindi di un problema di controllo ottimale.*

Si tratta di un problema di *controllo*. Agendo sui parametri di controllo – la forma del nuotatore nel nostro caso – vogliamo ottenere un obiettivo: spostare il nuotatore. Allo stesso modo guidare un'auto è un problema di controllo: agendo sul volante e l'acceleratore, cerchiamo di spostare il veicolo e portarlo in una data posizione. La teoria del controllo è l'insieme degli strumenti matematici utilizzati per dimostrare rigorosamente che un dispositivo è controllabile, vale a dire che il nostro nuotatore può effettivamente spostarsi effettuando certi movimenti. Questa teoria utilizza molti argomenti geometrici e la teoria delle equazioni differenziali ordinarie. Usando la teoria del controllo, possiamo dimostrare rigorosamente che alcuni nuotatori robot proposti dai fisici possono effettivamente spostarsi in un fluido in un regime dominato dalla viscosità. Le figure 3 e 4 mostrano esempi di tali robot.

In molti casi, sapere che è possibile controllare un sistema non è sufficiente. Vogliamo poterlo fare in modo ottimale rispetto ad un criterio, ad esempio, come quando viaggiando in auto si vuole ridurre al minimo il tempo di percorrenza o il carburante consumato. Siamo allora nel dominio matematico del controllo ottimale. Per tornare al nuoto, il problema di trovare uno stile che minimizzi l'energia meccanica spesa dal nuotatore per produrre un determinato spostamento (o quello di spostarsi più velocemente possibile con una data quantità di energia) è un problema di controllo ottimale.

È possibile costruire degli algoritmi per calcolare numericamente (utilizzando un computer) il movimento migliore per certi nuotatori. Dopo aver modellizzato il problema, si tratta di trovare una soluzione ad una equazione differenziale ordinaria a partire da un dato punto (la posizione e la forma del nuotatore iniziali) e arrivare ad un altro punto dato (la posizione e forma finali). Viene quindi utilizzato il cosiddetto "metodo dello sparo" (il cui nome, per analogia, proviene dal problema balistico di sparare una palla di cannone da un punto ad un altro, regolando la direzione del cannone). Per le equazioni differenziali necessarie per trovare il nuoto ottimale, conosciamo il punto di partenza e vogliamo mirare al punto d'arrivo regolando la velocità iniziale.

Queste ricerche teoriche del controllo ottimale sono molto attuali (si veda *Per saperne di più*). Chissà, ci permetteranno forse di progettare i micro-robot di domani!

Per saperne di più

Al lettore interessato all'approfondimento di questo argomento suggeriamo di leggere l'eccellente articolo

Purcell EM, (1977). *Life at low Reynolds number*, Am. J. Phys, 45(1):3–11.

Sono disponibili anche i lavori di ricerca di uno degli autori di questo articolo:

Alouges F., DeSimone A., Lefebvre A., (2008). *Optimal strokes for low Reynolds number swimmers: an example*, J. Nonlinear Sci., 18(3):277-302.

Alouges F., DeSimone A., Lefebvre A., (2009). *Biological fluid dynamics: swimming at low Reynolds numbers*, Encyclopedia of Complexity and System Science, Springer Verlag.

All'indirizzo http://www.cmap.polytechnique.fr/~alouges/nage.php si possono guardare dei video che mostrano le strategie ottimali di nuoto di certi meccanismi e il confronto con altri metodi meno efficaci. La figura 4 proviene da uno di essi.

Molto istruttivo è anche il filmato realizzato da Taylor G. I. Taylor (G. I. Taylor, *Low Reynolds Number Flow, National Committee for Fluid Mechanics Films*, Education Development Center. Inc., Newton, MA) - disponibile tra l'altro su YouTube - che mostra una gamma completa di particolarità dei flussi a basso numero di Reynolds.

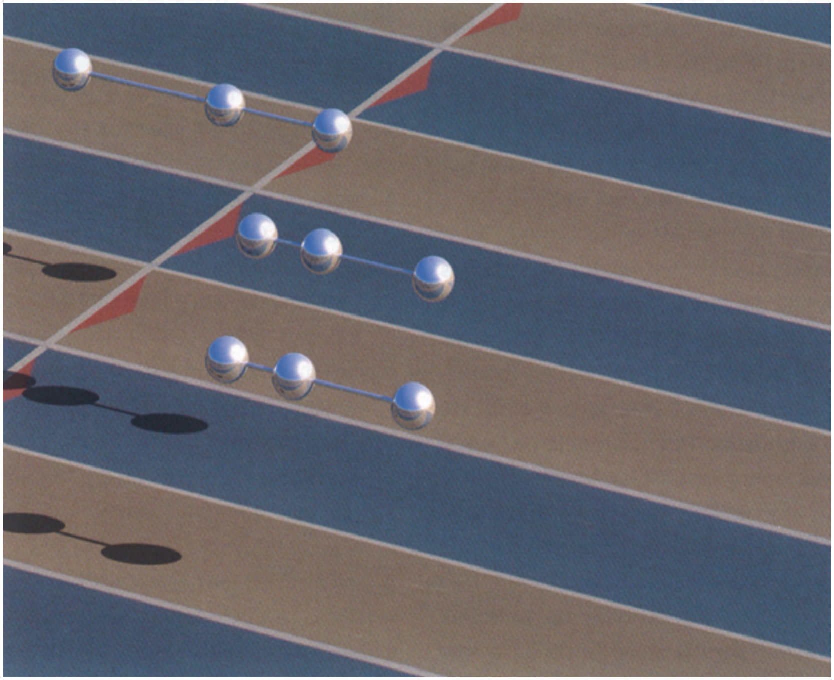


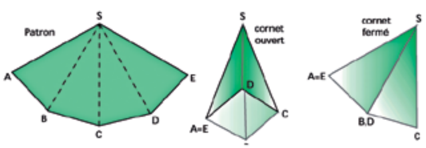
Figura 4. Simulazione numerica di una gara tra tre nuotatori composti da sfere allineate e collegate da arti allungabili. Il nuotatore in primo piano è stato calcolato in modo che il suo nuoto sia ottimale, e quindi si sposta più velocemente rispetto agli altri due a parità di energia. La gara si svolge da sinistra a destra.

Il teorema del soffietto

Étienne Ghys, *CNRS direttore di ricerca presso École Normale Supérieure de Lyon*

Un righello, una matita, del cartone, delle forbici e della colla: non c'è bisogno di altro per procurare ai matematici divertimento e problemi interessanti, il cui studio si rivela spesso, a posteriori e in modo del tutto inaspettato, di utilità nelle applicazioni.

Costruiamo una piramide di cartone... Per fare questo, cominciamo a tagliare un modello SABCDE da un foglio di cartone come indicato in figura, poi pieghiamo lungo le linee tratteggiate e infine incolliamo i lati AS ed ES.



Il risultato è una specie di cono il cui vertice è S e il bordo è il quadrilatero ABCD. Questo oggetto è *flessibile*, infatti il quadrilatero ABCD si può deformare aprendolo più o meno: la costruzione non è solidissima. Per completare la piramide bisogna tagliare ancora un quadrato di cartone ed incollarlo sul quadrilatero per formare la base. Dopo questa operazione la piramide risulta essere solida, rigida. Se si posa su un tavolo essa non crolla e se cerchiamo di deformarla con le mani (con non troppa forza) non vi riusciamo a meno di deformarne le facce di cartone. Analogamente, un cubo in cartone è *rigido* come tutti hanno potuto sperimentare. Che cosa succede ad un *poliedro* generico che abbia, ad esempio, un migliaio di facce? Il geode della Villette a Parigi è rigido? Quest'ultima domanda lascia intravedere come il tema della rigidità e della flessibilità non sia solamente teorico.

Il caso di poliedri convessi

Il problema della rigidità di oggetti di questo tipo è antichissimo. Euclide, con molta probabilità, ne era a conoscenza. Il grande matematico francese Adrien-Marie Legendre se ne è interessato verso la fine del XVIII secolo e ne ha parlato al suo collega Joseph-Louis Lagrange, il quale ha suggerito a sua volta al giovane Augustin-Louis Cauchy di studiare tale problema nel 1813. Sarà il primo risultato notevole del barone A.-L. Cauchy, che diventerà, poi, uno dei più grandi matematici del suo secolo.

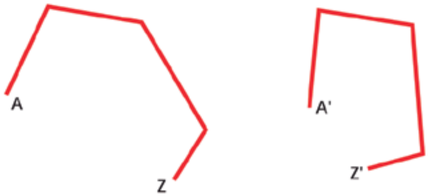
Augustin-Louis Cauchy (1789-1857),
uno dei grandi matematici del suo tempo.
(Archives de l'École Polytechnique)

Cauchy si è interessato a poliedri convessi, vale a dire poliedri che non hanno spigoli rientranti. Per esempio, la piramide che abbiamo costruito o il pallone da calcio sono convessi, mentre l'oggetto disegnato a destra nella figura qui sotto non lo è.

Il teorema formulato da Cauchy afferma che *tutti i poliedri convessi sono rigidi*. Questo significa che, se si costruisce un poliedro convesso con delle facce indeformabili (ad esempio in metallo) unite da cerniere lungo gli spigoli, la geometria globale dell'insieme impedisce alle giunture di muoversi. Il cono che abbiamo costruito è flessibile, ma questo non invalida il teorema: infatti, ad esso manca l'ultima faccia, che è quella che rende rigida la piramide.

Fare matematica significa *dimostrare* ciò che si afferma: la dimostrazione di Cauchy è superba (anche se alcuni hanno fatto notare, in seguito, che essa era incompleta). Purtroppo non è lo scopo di questo breve articolo quello di dare una idea della dimostrazione, ma mi piacerebbe evidenziarne un "lemma", cioè un passo della dimostrazione.

Posiamo a terra una catena formata da sbarre metalliche unite l'una all'altra per un estremo, come nella seguente figura:



Chiudendo gli angoli della catena, la distanza tra le estremità diminuisce: AZ è maggiore di A'Z'.

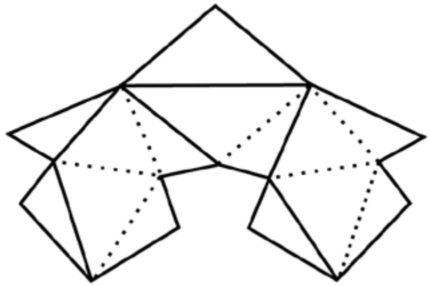
In ognuno degli angoli di questa poligonale muoviamo le sbarre consecutive in modo da diminuire l'angolo corrispondente; in questo modo accade che le due estremità della catena si avvicinano. Sembra ovvio! Provate a dimostrarlo...

Costruzione di un "flessidromo"

Per tanto tempo molti matematici si sono chiesti se anche i *poliedri non convessi* fossero rigidi; era possibile trovare una dimostrazione di rigidità che non usasse l'ipotesi di convessità? I matematici amano gli enunciati nei quali *tutte* le ipotesi sono utili per giungere alla tesi. Si è dovuto attendere più di 160 anni per rispondere a questo particolare interrogativo.

Nel 1977, il matematico canadese Robert Connelly ha provocato grande scalpore nella comunità matematica costruendo un poliedro (abbastanza complicato) che risultava essere flessibile e, per non contraddire il teorema di Cauchy, non convesso. In seguito la sua costruzione è stata un po' semplificata, in particolare da Klaus Steffen.

Nella figura seguente è rappresentato un modello che permette al lettore di costruire il cosiddetto flessidromo di Steffen. Tagliatelo e piegatelo lungo le linee; le linee continue sono gli spigoli sporgenti mentre le linee tratteggiate sono quelli rientranti. Incollate i bordi liberi nel modo più naturale ed otterrete una specie di *Shadok* che è veramente flessibile (almeno un po'...).



Il flessidromo di Steffen. I lati hanno lunghezza 5, 10, 11,12, e 17.

All'epoca i matematici furono incantati da questo nuovo oggetto. Fu costruito un modello metallico e posto nella sala da thè dell'Institut des Hautes Études Scientifiques a Bures-sur-Yvette vicino a Parigi e ci si poteva divertire a farlo muovere (per la verità non proprio bello e un po' cigolante). Si racconta che Dennis Sullivan ebbe l'idea di soffiare del fumo di sigaretta all'interno del flessidromo di Steffen e constatò che facendo muovere l'oggetto il fumo restava tutto all'interno... Da questo egli intuì che quando il flessidromo si deforma il suo volume rimane invariato. L'aneddoto è vero o è una leggenda? Comunque sia, Connelly e Sullivan congetturarono che *quando un poliedro si deforma il suo volume resta costante*. Non è difficile verificare questa proprietà nell'esempio particolare del flessidromo di Connelly o anche per quello costruito da Steffen (a prezzo di calcoli complicati, ma privi di interesse matematico): tuttavia la congettura in questione considera **tutti** i poliedri, compresi quelli che non sono mai stati praticamente costruiti! Tale problema è stato chiamato "congettura del soffietto": il soffietto del focolare emette aria quando viene compresso, in altri termini il suo volume diminuisce (del resto proprio questa è la sua funzione). Sicuramente un soffietto reale non corrisponde alla definizione di poliedro, in quanto esso è in pelle e le sue facce si deformano costantemente, contrariamente a quanto accade per i poliedri le cui facce sono rigide.

Nel 1997 Connelly e due altri matematici I. Sabitov e A. Walz sono finalmente riusciti a dimostrare questa congettura. La loro dimostrazione è grandiosa e mostra ancora una volta le interazioni tra tutte le parti della matematica: per risolvere questo problema prettamente geometrico, gli autori hanno utilizzato metodi molto raffinati di moderna algebra astratta. Non si tratta di una dimostrazione che Cauchy "avrebbe potuto trovare": le tecniche matematiche dell'epoca erano insufficienti.

Vorrei ricordare una formula che a volte si impara alla scuola secondaria. Se le lunghezze dei lati di un triangolo sono a, b e c, si può calcolarne facilmente la superficie. Per fare ciò si calcola a parte il *semi-perimetro* $p=(a+b+c)/2$ e poi si ottiene l'area S estraendo la radice quadrata di $p(p-a)(p-b)(p-c)$. Questa elegante formula porta il nome del matematico greco Erone e proviene dalla notte dei tempi. Si può calcolare con una formula analoga il volume di un poliedro conoscendone la lunghezza degli spigoli? I tre matematici sopracitati hanno dimostrato che è possibile fare ciò.

Essi cominciarono con un poliedro costruito a partire da un modello formato da un certo numero di triangoli chiamando l_1, l_2, l_3, … le lunghezze dei lati di questi triangoli (eventualmente molto numerosi). Essi provarono quindi che il volume V del poliedro deve soddisfare una certa equazione di grado n-esimo, cioè della forma $a_0 + a_1V + a_2V^2 + \ldots + a_nV^n = 0$.

La cupola della Villette (1730 triangoli e 36 metri di diametro). Accanto una cupola più piccola.

Il grado n dipende dal modello utilizzato ed i coefficienti dell'equazione a_0, a_1, ... dipendono esplicitamente dalla lunghezza dei lati l_1, l_2, l_3, ... In altri termini, se si conosce il modello e le lunghezze dei lati, è possibile trovare l'equazione. Se il lettore ricorda che, in generale, un'equazione ha una soluzione quando è di grado 1, due quando il grado è 2, può indovinare che un'equazione di grado n non ha più di n soluzioni. Quindi, in conclusione, se conosciamo il modello e le lunghezze dei lati di un poliedro, pur non conoscendo l'esatto valore del volume, sappiamo che tale volume non può assumere che un numero finito di valori... Quando il flessidromo si deforma, il suo volume non può, quindi, variare (altrimenti il volume avrebbe un'infinità di valori successivi); questo volume è "bloccato" e la congettura del soffietto è dimostrata...

Il classico, il semplice e l'applicato

Questo problema è utile? È interessante? Quando un problema matematico è interessante? Questa è una domanda difficile sulla quale i matematici, sicuramente, riflettono da molto tempo. Ecco qualche possibile risposta, qualche tentativo di dare un "indice di qualità". L'antichità è un primo criterio: i matematici sono molto sensibili alla tradizione, ai problemi enunciati da molto tempo, sui quali si sono interrogati molteplici generazioni di studiosi. Un buon problema, inoltre, deve avere un enunciato semplice, la cui dimostrazione porti a sviluppi sorprendenti, se è possibile, mettendo in relazione aree molto diverse della matematica. Il problema della rigidità di cui stiamo parlando è interessante da entrambi questi punti di vista. Capire se un problema matematico abbia applicazioni nella vita pratica è più sottile. I matematici rispondono in varie maniere. Senza dubbio le domande "pratiche", provenienti per esempio dalla fisica, servono molto spesso come motivazione alla matematica. A volte, cercando di risolvere un problema molto concreto, il matematico generalizza il problema alla risoluzione di un modello astratto: egli si serve del problema iniziale come *sorgente di ispirazione* e l'effettiva risoluzione del problema concreto non è più la vera motivazione. Il problema della rigidità appartiene a quest'ultima categoria; infatti, la sua origine dalla fisica è abbastanza chiara: la stabilità e la rigidità di strutture, per esempio, metalliche. Al momento, i teoremi e le costruzioni di Connelly non sono di nessuna utilità per gli ingegneri. Appare evidente, tuttavia, che questo tipo di ricerca non mancherà, in avvenire, di portare ad una migliore comprensione globale della rigidità di svariate strutture costituite da un gran numero di elementi individuali (macromolecole, edifici, ecc...). Si tratta, dunque, di ricerche teoriche e "disinteressate", ma che hanno buone possibilità di essere utili nel futuro.

> *Questo tipo di ricerca non mancherà, in avvenire, di portare ad una migliore comprensione globale della rigidità di svariate strutture costituite da un gran numero di elementi individuali.*

Per saperne di più

Berger M., (1977). Géométrie, v. 3. – *Convexes et polytopes, polyèdres réguliers, aires et volumes* (CEDIC/Nathan Information).

Connelly R., Sabitov I., Walz A., (1997). *The bellows conjecture*, Beiträge Algebra Geom., 38, n° 1, p. 1-10.

Connelly R., (1977). *A counterexample to the rigidity conjecture for polyhedra*, Institut des Hautes Études Scientifiques, Publication Mathématique n° 47, p. 333-338.

Kuiper N. H., (1979). *Sphères polyédriques flexibles dans* E^3, d'après Robert Connelly, Séminaire Bourbaki, 30 e année (1977/78), exposé n° 514, pp. 147-168 (Lecture Notes in Math. 710, Springer).

Riconoscere lo Spam: un gioco da ragazzi?

Tristan Mary-Huard, *ricercatore INRA presso INRA-AgroParisTech*

Come distinguere automaticamente lo spam da un messaggio normale. I filtri anti-spam analizzano i messaggi di testo utilizzando algoritmi di classificazione a forma di alberi. Questi comprendono un numero ottimale di nodi corrispondenti ad altrettante domande pertinenti per determinare la natura di un messaggio.

Alcuni anni fa, un famoso produttore di giocattoli ha proposto il gioco "*Indovina Chi?*". Il principio del gioco è semplice: ogni giocatore deve trovare tra una collezione di personaggi quello scelto dal suo avversario. Per far questo, ogni giocatore a turno pone una domanda del tipo: "Il personaggio ha i baffi?", oppure: "Il personaggio ha gli occhiali?". Il primo giocatore che riesce a trovare il personaggio dell'avversario vince la partita. La vittoria, quindi, va al giocatore che pone la sequenza di domande più pertinenti per identificare il più rapidamente possibile il personaggio misterioso.

La storia non dice se Leo Breiman, matematico e ricercatore americano, e i suoi colleghi presso l'Università di Berkeley fossero stati giocatori appassionati di "*Indovina chi?*". Tuttavia il metodo da lui proposto nel 1984 segue strettamente la logica di quel gioco e lo utilizza per risolvere i problemi di classificazione. Questo metodo è stato più volte sfruttato e serve oggigiorno, tra l'altro, alla rilevazione dello spam.

Tra tutti i messaggi ricevuti da un individuo per posta elettronica, ci sono due tipi di messaggi: "regolari" (inviati da amici o da siti internet ai quali si è abbonati) e messaggi "spam" (comunicazioni elettroniche non richieste). Questi messaggi di spam vengono solitamente inviati a migliaia, anche milioni di persone sulla rete, e le conseguenze di tali invii di massa non sono trascurabili: in Francia, si stima che il 95% delle mail ricevute nel 2009 fossero messaggi di spam.

Una possibilità è quella di classificare i messaggi mediante lo studio della frequenza con cui certe parole occorrono.

Per evitare che gli utenti trovino la loro cartella di posta intasata, i fornitori di servizi internet cercano di sviluppare metodi di filtraggio anti-spam in grado di distinguere automaticamente un messaggio di spam da uno regolare. Tali metodi sono chiamati algoritmi di classificazione.

"Vincere", "lotteria": Le parole tipiche dello spam

Su quali informazioni può basarsi l'algoritmo di classificazione? Ogni messaggio contiene del testo. È dunque a partire dall'analisi del contenuto del testo che bisogna classificare un messaggio come spam o regolare. Spesso lo spam cattura l'attenzione del destinatario con la promessa di un guadagno (in genere economico) e chiede all'utente di collegarsi ad un sito che in questo modo diventerà molto frequentato. Questi messaggi sono quindi caratterizzati dalla presenza di parole come "vincere", "lotteria" o "cliccare" nel corpo del testo. Ci sono quindi delle parole che possono rivelare la natura del messaggio. Quindi, una possibilità è quella di classificare i messaggi dallo studio della frequenza con cui certe parole occorrono.

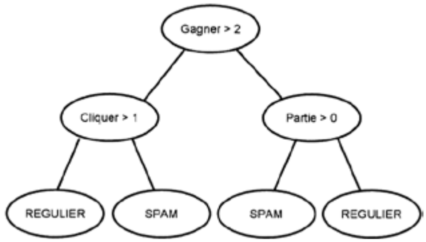


Figura 1. Esempio di albero di classificazione per classificare i messaggi di posta elettronica.

L'algoritmo proposto da Breiman e i suoi colleghi è illustrato in figura 1. Rappresenta il percorso che deve seguire un messaggio per essere classificato come spam o regolare. Il percorso inizia nella parte superiore, leggendo la domanda scritta nel nodo radice (le ellissi sono chiamati *nodi*, le linee tra due ellissi *rami*, e il primo nodo superiore è chiamato il *nodo radice*). Se la risposta alla domanda è sì, vale a dire se la frequenza di occorrenza della parola "vincere" è strettamente maggiore di 2, il messaggio continua il suo percorso attraverso il ramo di destra. Altrimenti, il messaggio prosegue nel ramo sinistro. Il messaggio raggiunge quindi un altro nodo, e come prima continua il suo percorso verso destra o sinistra a seconda della risposta alla nuova domanda. Questo viene ripetuto finché il messaggio raggiunge una foglia, vale a dire un nodo dal quale non parte nessun ramo. Nella foglia raggiunta si legge come il messaggio viene classificato. Si noti che il vocabolario utilizzato (radice, nodi, rami e foglie) non è dovuta al caso: l'algoritmo di classificazione di Breiman è chiamato infatti albero di classificazione.

Consideriamo i seguenti messaggi

Messaggio 1

Congratulazioni,

hai appena vinto la nostra grande lotteria su internet.

Clicca qui per ricevere 1.000.000 dollari!!!

E per vincere altri premi, clicca qui.

Messaggio 2

Ciao Benoît,

domani gioco contro Bruno.

Se vinco la partita mi qualifico direttamente.

Se non vinco domani, ma vinco la prossima, passo in corsia 3 l'anno prossimo!

Abbracci, Sandra

Nel primo messaggio, la parola "vittoria" appare solo due volte. La risposta alla domanda del nodo principale è negativa, e il messaggio segue il ramo di sinistra. Arrivando al nodo successivo, vediamo che la parola "clicca" appare due volte. Il messaggio prende quindi il ramo di destra, raggiungendo alla fine una foglia SPAM. Il messaggio verrà quindi classificato come spam. Nel secondo messaggio, la parola "vincere" appare tre volte, ma è qui usata per indicare "vincere una partita" piuttosto che "fare soldi". Questo uso alternativo della parola "vincere" è tenuto in conto (dal nodo situato in basso a destra del nodo radice) e il messaggio sarà classificato REGOLARE dall'albero della figura 1.

Naturalmente, l'albero e i messaggi qui considerati sono solo un esempio molto semplificato del problema generale. Gli alberi di classificazione utilizzati nella pratica sono spesso più grandi (in termini del numero di nodi), e le domande connesse ad ogni nodo riguardano la presenza di decine di parole diverse.

Valutare i diversi alberi

Rimane la questione della scelta dell'albero di classificazione. Come si sceglie il numero di domande da porre? O su quali parole devono vertere le diverse domande? Queste scelte possono essere effettuate nel modo seguente: ciascuno può proporre l'albero di classificazione che gli sembra più efficace e tutti questi alberi possono essere misurati su di un campione di prova. Un campione di prova è un insieme di messaggi la cui natura è nota, perché sono già stati letti e classificati come spam o regolare da parte di una persona.
Gli alberi di classificazione vengono applicati uno dopo l'altro al campione di prova e per ogni messaggio il risultato dell'algoritmo di classificazione viene confrontato con la vera classificazione. Un albero di classificazione che troppo spesso classifica come normale messaggio di spam o viceversa viene escluso. Tra gli alberi rimanenti, verrà scelto quello con il minor numero di nodi, vale a dire quello che permette di stabilire la natura del mes-

saggio con il minor numero di domande. Il vincitore è quindi quell'albero che pone le domande più rilevanti per determinare il più rapidamente possibile la natura dei messaggi. Ritroviamo quindi la strategia del gioco *Indovina Chi?*!

Migliorare il filtraggio

La longevità degli alberi di classificazione, utilizzati ormai da 25 anni, e la diversità delle applicazioni che hanno trovato (ad esempio, nei campi della biologia, dell'ecologia e della medicina, si veda il riquadro) è dovuto non solo al fatto che le prestazioni ottenute in pratica da questo algoritmo sono soddisfacenti. In effetti, dal punto di vista matematico, l'algoritmo solleva molte domande: si può utilizzare per qualsiasi problema di classificazione, o ci sono problemi per i quali l'algoritmo potrebbe rivelarsi inefficace? Possiamo dimostrare che gli alberi di classificazione sono migliori (in termini di prestazioni di classificazione) di altri algoritmi comunemente utilizzati? Miglioreremmo le prestazioni utilizzando più alberi di classificazione, invece di uno, come quando si convoca una giuria di esperti, invece che uno solo? Molti articoli sono stati dedicati a queste domande, e più in generale a determinare le proprietà degli alberi di classificazione, articoli che hanno contribuito a giustificare teoricamente il loro utilizzo. Ma questa ricerca matematica sugli alberi di classificazione, lungi dal servire unicamente a convalidare il metodo iniziale, ha permesso anche ai ricercatori di suggerire come modificare l'algoritmo e migliorarne le prestazioni. Ecco quindi, l'applicazione degli alberi di classificazione per filtrare lo spam, che si è considerevolmente affinato nel corso degli anni. Da un unico albero di classificazione, siamo passati ad una foresta (metodo delle *foreste random*, vedi riquadro) e gli alberi sono diventati dinamici: mano a mano che il proprietario della posta elettronica riceve dei messaggi ciascun albero continua a raccogliere informazioni sulle frequenze di parole per migliorare le sue prestazioni di classificazione. Gli alberi di classificazione, e in generale tutti gli algoritmi di classificazione, costituiscono quindi un tema di ricerca molto attivo, al quale sono dedicate numerose e prestigiose conferenze internazionali di matematica.

Mano a mano che il proprietario della posta elettronica riceve dei messaggi, ciascun albero continua a raccogliere informazioni sulle frequenze delle parole per migliorare le sue prestazioni di classificazione.

D'altra parte gli spammer non stanno a guardare e i messaggi di spam sono oggigiorno scritti in modo tale da aggirare i filtri. Le parole che potrebbero tradire la natura del messaggio sono accuratamente evitate, o intenzionalmente scritte con errori ortografici per renderle più difficili da identificare da parte dell'algoritmo di classificazione. La gara tra gli spammer e i filtri è appena iniziata e continuerà anche in futuro a porre ai matematici nuovi problemi appassionanti ... e altre critiche verso lo sviluppo informatico ed economico della nostra società!

Il metodo delle foreste random

di Robin Genuer, *ricercatore presso Université di Bordeaux Segalen*

Il metodo delle *foreste random* (foreste aleatorie) può essere molto utile nelle applicazioni della medicina. Ad esempio, quando si studiano i fattori di rischio di una malattia e non si dispone di informazioni sui geni dei soggetti affetti o no da questa malattia. L'obiettivo è quindi di trovare i geni che permettono di distinguere con maggiore probabilità i soggetti malati da quelli sani.

Tuttavia, nella maggior parte dei casi, abbiamo informazioni su decine di migliaia di geni, e quindi di altrettante potenziali domande per biforcare un nodo dell'albero. Ciò rende molto difficile offrire un buon albero "al primo colpo" (fare una partita di *Indovina chi?* con 10.000 domande possibili sembra davvero molto arduo!).

L'idea delle foreste random è quello di costruire una collezione alberi dove per ciascun albero si restringe il numero di domande da fare. Dopo aver messo insieme tutti questi alberi, si può valutare la classificazione proposta dalla foresta (su di un campione di prova).

Molti studi indicano che una foresta sembra comportarsi meglio di un solo albero per questo tipo di problemi (anche se le ragioni di questo fenomeno sono ancora relativamente sconosciute).

Per saperne di più

Breiman L., Friedman J., Olshen R., Stone C., (1984) *Classification And Regression Trees*, Chapman & Hall.

Articolo sullo Spam di Wikipedia, (http://it.wikipedia.org/wiki/Spam).

European Network and information Security Agency (ENISA), *Spam Survey*, 16 décembre 2009

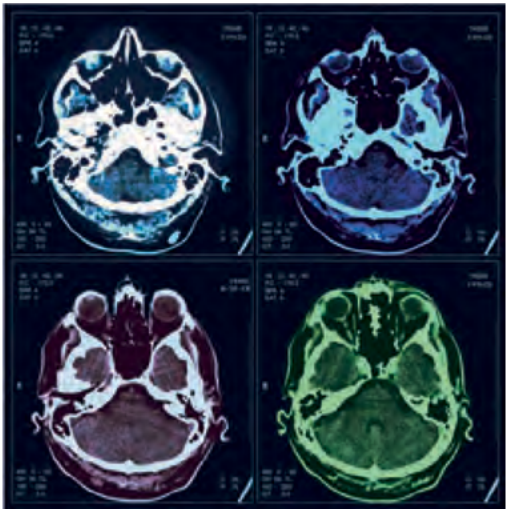
Surhone LM, Tennoe MT, Henssonow SF, (2010) *Random Forest*, Verlag.

L'arte di tagliare teste senza ferire

Erwan Le Pennec, *ricercatore presso INRIA all'Università Paris-Sud*

Il principio dalla tomografia è quello di ricostruire un oggetto da una raccolta di sue radiografie. Questo è noto in matematica come un problema inverso. La risoluzione è difficile e sempre fonte di nuove domande per matematici.

Sembrerebbe difficile "tagliare" una testa per osservare il suo interno senza far male al suo proprietario. Eppure questo è quello che le équipe mediche possono ottenere utilizzando la TC (Tomografia Computerizzata). Questo strumento diagnostico è disponibile grazie ai lavori di Godfrey Hounsfield e Allan McLeod Cormack, due fisici che hanno condiviso il premio Nobel nel 1979 per questa invenzione. La TC non potrebbe esistere senza il matrimonio di fisica, matematica e informatica.



Problema diretto, problema inverso

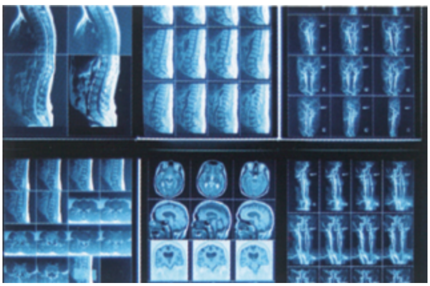
Diversamente dalla radiografia convenzionale a raggi X, l'immagine di una TC non è ottenuta mediante una misurazione fisica, ma è creata da un computer utilizzando algoritmi matematici a partire da una collezione di radiografie della testa del paziente prese da angolazioni diverse. Per riuscire in una tale impresa, è necessario risolvere due problemi: spiegare come questo insieme di radiografie dipende dall'interno della testa del paziente (questo è chiamato *problema diretto*) e capire come mettere insieme la raccolta delle radiografie per ricostruire cosa accade all'interno della testa del paziente (parliamo qui di *problema inverso*).

Il problema diretto è lo stesso di quello delle radiografie a raggi X. Il principio è stato scoperto da Wilhelm Conrad Roentgen nel 1895 ed è stato immediatamente utilizzato per diagnosticare lesioni interne. È stato ampiamente studiato da un punto di vista fisico: è ben noto che l'immagine radiografica dipende principalmente dalla quantità di materiale assorbente attraversato dai raggi X tra la sorgente ed il sensore. Questo è il motivo per cui le ossa, più assorbenti dei tessuti, sono le strutture più visibili. Tenendo conto delle caratteristiche fisiche degli strumenti a raggi X utilizzati, è dunque facile prevedere matematicamente, in modo molto preciso, la radiografia di un oggetto che sia completamente noto. Lo stesso ovviamente vale per una raccolta di radiografie prodotte da uno scanner per la TC.

Sebbene molto precisa, questa modellizzazione non è perfetta: la radiografia misurata sarà leggermente diversa da quella attesa a causa delle semplificazioni fisiche, dell'imprecisione di vari sensori e di altri fattori.

Ricostruire l'oggetto a partire dalle radiografie

Ricostruire l'oggetto a partire dalla collezione di radiografie, il problema inverso, risulta essere un problema molto più difficile. Dobbiamo riuscire ad invertire la procedura di acquisizione della scansione mediante una procedura implementabile con un computer. Inoltre, è necessario che le piccole imprecisioni nel modello non inducano grandi errori sul risultato. È la messa a punto di una soluzione a questo problema che ha permesso la nascita della prima generazione di apparecchi per la TC alla fine degl'anni '60. Da allora, gli sviluppi tecnologici delle sorgenti di raggi X, dei sensori nonché di altri apparati strumentali hanno migliorato la qualità delle immagini ottenute. Questa sarebbe oggi molto più scarsa senza il lavoro dei matematici e lo sviluppo di nuovi algoritmi di inversione.



Un principio semplice, spesso chiamato il principio del *rasoio di Ockham*, è la fonte della maggior parte dei metodi moderni per risolvere questo problema. Guglielmo di Ockham, un monaco francescano e filosofo del XIV secolo, ha spiegato che “le ipotesi non devono essere assunte senza necessità”. Nella sua formulazione moderna, il rasoio di Ockham è un principio di parsimonia: una buona soluzione è una soluzione *coerente* con i dati osservati senza essere inutilmente *complessa*.

Per applicare questo principio di buon senso, dobbiamo ovviamente fornire una definizione precisa (e matematica) di “coerenza” e di “complessità”, nonché proporre un metodo per ottenere una soluzione che sia coerente e poco complessa.

Una buona soluzione è una soluzione coerente con i dati osservati senza essere inutilmente complessa

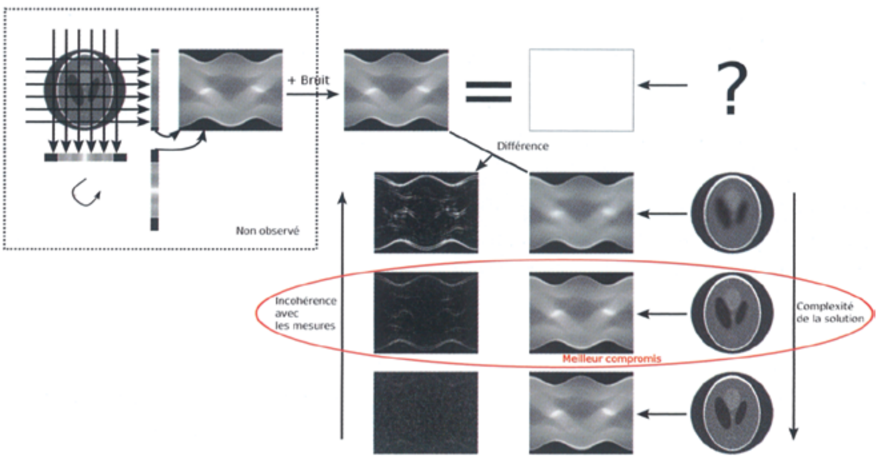
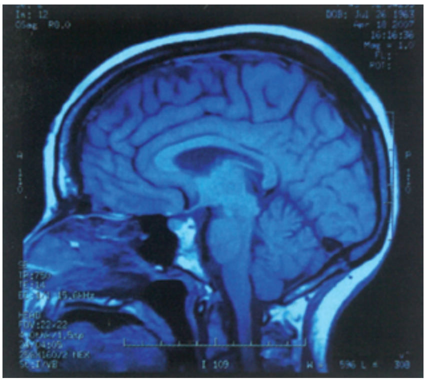


Figura 1. A sinistra si trova un’immagine test chiamata fantoccio di Shepp-Logan, spesso utilizzato per la valutazione dei metodi tomografici; la foto immediatamente alla sua destra corrisponde alle misure di uno scanner perfetto (questa immagine bidimensionale si ottiene giustapponendo le “fette” per i diversi piani di taglio). L’immagine seguente corrisponde alle misure leggermente disturbate; infine sotto sono rappresentati tre equilibri differenti tra l’incoerenza (l’immagine della colonna di sinistra è tanto più chiara quanto più è grande la discrepanza con le misure) e la complessità (l’immagine della colonna di destra corrisponde a delle soluzioni di complessità crescente dall’alto verso il basso).

Nessuno dei metodi per risolvere il problema raccoglie l'unanimità dei consensi: i matematici, così come i produttori di apparecchi per la TC, continuano ad esplorare altre possibilità. Uno dei metodi sul quale ha lavorato l'autore del presente articolo, è illustrato in figura 1.

Si è osservato che la soluzione *migliore* sembra essere quella che realizza il *miglior* compromesso tra coerenza e complessità.



Risonanza magnetica, indagini sismiche e problemi connessi

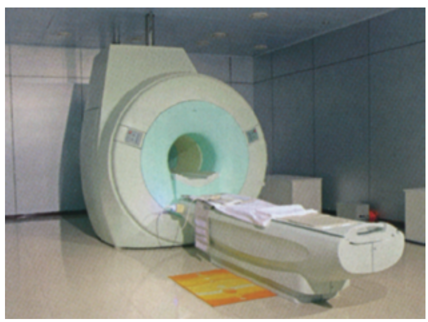
Esistono molti problemi simili nell'elaborazione delle immagini che includono, per esempio:

– La risonanza magnetica nucleare (RMN), in cui si cerca di determinare le strutture all'interno di un corpo misurando la loro interazione con i campi magnetici.

– La *deconvoluzione ottica*, dove si cerca di eliminare la sfocatura causata dalla lente in una fotografia.

– Le immagini sismiche, dove l'obiettivo è quello di individuare le strutture del sottosuolo a partire dai rilevamenti della propagazioni delle onde.

In ciascuna di esse, è possibile identificare un problema diretto, ovvero, lo studio di come le misure dipendono da un oggetto, e un problema inverso, nel quale si cerca di risalire all'oggetto dalle misure. Per ciascuno di questi problemi, il principio del rasoio di Ockham si è dimostrato essere efficace per ottenere risultati sia teorici che pratici. Le questioni intorno a questi problemi inversi sono ben lungi dall'essere risolte: nuovi dispositivi d'acquisizione vengono regolarmente introdotti e nuovi strumenti di calcolo sempre più potenti consentono di considerare metodi di complessità crescente, così come nuovi strumenti matematici consentono di esplorare nuovi metodi …



Per approfondire ...

La modellizzazione più comune del problema menzionato in questo articolo è l'osservazione di un oggetto O attraverso una collezione di sue radiografie R ottenute mediante l'applicazione di una funzione Φ proveniente dal modello diretto al quale si aggiunge una perturbazione B sconosciuta. Si ha quindi

$$R = \Phi(O) + B$$

e desideriamo trovare nel miglior modo O a partire dalla collezione delle radiografie R.

Un'idea naturale per un matematico è determinare l'inversa Φ^{-1} della funzione Φ e applicarla ad R per ritrovare O. Sfortunatamente, ci si accorge che l'oggetto così ottenuto

$$\Phi^{-1}(R) = \Phi^{-1}(\Phi(O) + B)$$

può essere molto diverso da O anche se B è piccolo. Si dice che l'inversione è *instabile*.

Per applicare il principio di Ockham, è necessario definire il criterio di *coerenza* così come quello di *complessità*.

Il matematico definisce una funzione d'incoerenza $F(R, O)$ che quantifica la differenza tra la radiografia osservata R e quella prevista dal modello diretto di un oggetto O: questa funzione è piccola quando l'oggetto O è coerente con le osservazioni e grande quando è incoerente.

Allo stesso modo, si definisce una misura di complessità $C(0)$, che sarà grande quando l'oggetto è complesso e piccola quando l'oggetto è semplice.

È possibile, quindi, interpretare il principio di Ockham come la ricerca di un oggetto O che realizzi il miglior compromesso tra queste due grandezze, cioè un oggetto per cui la somma

$$F(R,O) + C(O)$$

sia la più piccola possibile.

Gli esperimenti numerici della figura sono stati ottenuti scegliendo come criterio di coerenza essenzialmente la distanza usuale tra l'immagine dell'oggetto $\Phi(O)$ e R, e per criterio di complessità, la parsimonia, il numero di coordinate non nulle in una rappresentazione adeguata. Questa scelta offre delle garanzie teoriche sulla ricostruzione e un algoritmo efficiente per determinare il miglior compromesso.

Climatologia e statistica

Yiou Pascal, *ricercatore presso CEA*
Philippe Naveau, *ricercatore CNRS presso Institut Pierre Simon Laplace*

Lo studio del clima e delle sue variazioni si basa su un gran numero di strumenti e concetti statistici. Ciò si riflette nel linguaggio utilizzato dai vari rapporti dell'IPCC (Comitato internazionale di esperti sui cambiamenti climatici), che mettono un forte accento sulle incertezze e la loro quantificazione.

La stima del cambiamento climatico nel corso degli ultimi secoli e la visione prospettica del comportamento degli ultimi decenni in questo contesto plurimillenario, è una delle principali sfide scientifiche attuali. Osservando i grafici delle temperature medie rilevate sul globo o su ciascun emisfero (o in alcune sottoregioni), si constata un aumento complessivo delle temperature nel corso del XX secolo, con un'accelerazione a partire dagli anni 1970. Un gran numero di argomentazioni, nonché di misurazioni fisiche e termodinamiche, dimostrano che questo aumento è dovuto alle attività umane, responsabile del rilascio dei gas a effetto serra nell'atmosfera.

Clima e meteorologia a confronto

Il primo ingresso della statistica nella climatologia è rintracciabile nella definizione stessa di clima e nella differenza tra i termini meteorologia e climatologia. Per citare un celebre aforisma attribuito a E.N. Lorenz, il meteorologo che ha inventato l'espressione "effetto farfalla" e che ha introdotto lo studio del caos nelle scienze dell'atmosfera, "climate is what you expect, meteorology is what you get", ovvero, "il clima è quello che ti aspetti, la meteorologia è quello che succede realmente". La climatologia e la meteorologia usano e descrivono le stesse variabili fisiche (temperatura, pressione, velocità del vento, precipitazioni ...), ma i meteorologi sono interessati allo stato di queste variabili in un determinato luogo e in un dato momento, mentre i climatologi osservano i valori medi in una stagione o in una regione. Questi valori medi possono essere le variabili fisiche abituali (per esempio, la temperatura media invernale nell'Île-de-France) o altre grandezze più sofisticate, come la varianza della velocità del vento o l'insolazione media in una regione. In generale, le variabili climatiche sono definite per periodi di 30 anni (ad esempio 1961-1990) per motivi euristici riferiti all'inferenza statistica delle grandezze medie. È per questa ragione che non è ragionevole confrontare le tendenze di temperatura per un periodo di dieci anni o meno, con le tendenze stimate su tutto il XX secolo.

Il primo ingresso della statistica nella climatologia si incontra nella definizione stessa di clima.

L'evoluzione di temperature medie

Una domanda scientifica legittima è se i valori delle temperature rilevati attualmente siano già stati riscontrati nell'ultimo millennio, dato che la configurazione dei continenti e la posizione della Terra intorno al Sole non sono praticamente cambiati. Per affrontare questo compito, dato che non ci sono registrazioni termometriche prima del XVII secolo, si utilizzano degli indicatori climatici come gli anelli degli alberi, i sedimenti lacustri, i carotaggi dei ghiacciai, o anche delle testimonianze scritte (come le date dei raccolti). La combinazione di queste registrazioni indirette si realizza attraverso tecniche statistiche multivariate a volte sofisticate (come, ad esempio, le *wavelet*, i *modelli bayesiani gerarchici*, le *componenti principali*, ...), tecniche che permettono di ricostruire un segnale interpretato come un cambiamento di temperatura durante gli ultimi secoli. Il risultato che più colpisce di questi studi, realizzati da una quindicina di gruppi di ricerca in tutto il mondo, è la netta variazione delle temperature. Il grafico che ne risulta evidenzia una forma chiamata a "bastone da hockey": si può osservare una lieve diminuzione della temperatura osservata tra l'anno Mille e il 1900 e un rapido aumento nel corso del XX secolo (figura 1). Questo bastone da hockey è un leggermente storto, con un clima caldo (tra il 1000 e il 1400) e una piccola era glaciale fredda (tra il 1500 e il 1800). L'utilizzo di metodi statistici consente di ottenere intervalli di confidenza adeguati per rispondere alla domanda: "Qual è la probabilità che la temperatura attuale fosse parte della gamma di temperatura ricostruite a partire dall'anno mille?" La risposta è: una probabilità estremamente bassa, indi-

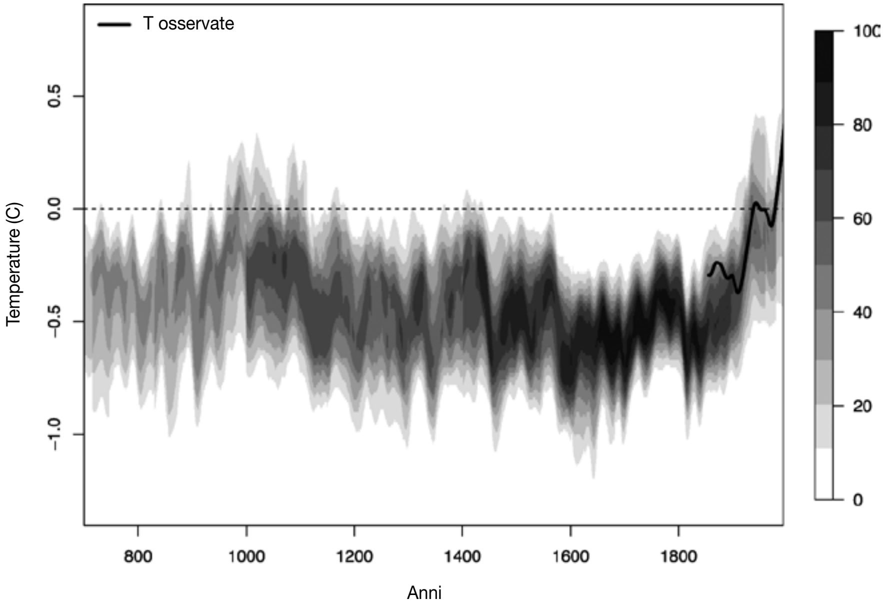


Figura 1. Ricostruzione della temperatura dell'emisfero settentrionale a partire dall'anno 800. Le temperature sono espresse come differenza dalla media del periodo 1961-1990 (linea tratteggiata). I livelli di grigio indicano la probabilità (in %) che la temperatura ricostruita per un dato anno si trovi in una certa gamma di temperature. La linea continua indica le temperature medie nell'emisfero settentrionale (adattato dal rapporto IPCC 2007).

pendentemente dal metodo statistico utilizzato nella ricostruzione! In altre parole, non vi è alcuna possibilità che il cambiamento climatico attuale sia parte di una normale evoluzione del clima.

Qual è la probabilità che la temperatura attuale fosse parte della gamma di temperature ricostruite a partire dall'anno Mille?

Variazioni intorno al valore medio

Una conseguenza dello studio delle variazioni delle temperature medie riguarda le variazioni rispetto al valore medio: possiamo dire che si osservano sempre di più degli eventi estremi (come ondate di calore anomale, precipitazioni pesanti, ecc.) alla fine del ventesimo secolo?

Per rispondere a questa domanda, i climatologi possono utilizzare modelli statistici costruiti ad hoc. Questo approccio richiede la disponibilità di osservazioni di qualità e durata sufficienti.

Se siamo interessati alle ondate di calore, al freddo o alla pioggia intensa, il quadro statistico che pare più appropriato è quello della Teoria dei Valori Estremi, elaborato da Emile Gumbel, nel 1940. Questa teoria descrive la distribuzione dei valori grandi di una variabile aleatoria continua X. La sfida principale è quella di descrivere i comportamenti estremi di una legge che non è nota a priori (che è il caso della maggior parte delle variabili che si misurano).

Teoria dei valori estremi

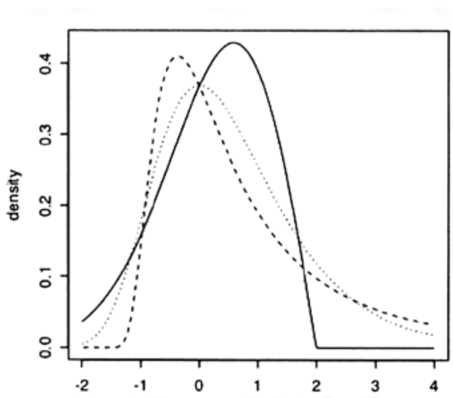


Figura 2. Densità di probabiltà della legge di Fréchet (--), Weibull (-) e Gumbel (..).

Massimo di n variabili aleatorie → **Distribuzione di Fréchet, Weibull o Gumbel**

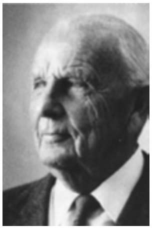
n tende ad infinito

M-R. Fréchet

I matematici degl'anni 70 hanno risolto questo problema affermando che se la distribuzione dei massimi di X converge ad una distribuzione, allora questa è una delle tre distribuzioni estreme: le distribuzioni di Gumbel, Fréchet e Weibull (che differiscono essenzialmente nel modo in cui la loro densità si comporta per valori molto grandi). Per tornare al clima, ci interessa il tipo di distribuzione di probabilità che seguono le temperature massime per ogni stagione. Nella maggior parte delle regioni, le distribuzioni delle temperature massime seguono la distribuzione di Weibull, vale a dire che la distribuzione dei massimi tende a 0 più velocemente di una di una distribuzione esponenziale (o gaussiana). Tuttavia, le leggi della massima precipitazione seguono le distribuzioni di Gumbel o di Fréchet: i valori possono essere molto grandi in maniera relativamente frequente.

E-J. Gumbel

L'ultimo rapporto speciale del IPCC sugli eventi estremi mostra che il numero di ondate di calore in Europa è cresciuto nel XX secolo. Questo vale anche per quello delle precipitazioni intense nel sud dell'Europa, anche se questa regione vede un aumento del numero dei giorni senza pioggia.



W. Weibull

Il modo in cui le proprietà di queste distribuzioni estreme cambiano nel tempo sono degli argomenti di ricerca molto attivi nel campo della climatologia statistica. Una stima di questi estremi e dei loro margini di errore è essenziale per fare previsioni climatiche che siano utili a vari settori economici, come quello dell'energia, dell'agricoltura, dell'assicurazioni e dei trasporti.

Per saperne di più

Jeandel C., Mosséry R., (éd.) (2011). *Le Climat à Découvert : Outils et méthodes en recherche climatique*, CNRS Éditions, Paris.

Cercare nel web: nient'altro che un punto fisso e qualche algoritmo

Serge Abiteboul, *direttore di ricerca Inria presso École Normale Supérieure de Cachan*

Il Web ci offre una notevole quantità di informazioni, diverse decine di miliardi di documenti. Senza i motori di ricerca, questi sistemi sempre più sofisticati che ci aiutano a focalizzarci su di un piccolo numero di pagine, il Web sarebbe una massa di rifiuti, enorme e inutilizzabile. Il ruolo di questi sistemi è quello di far emergere dalla comunità degli internauti una intelligenza collettiva in grado di valutare, classificare, filtrare le informazioni. Come fanno i motori di ricerca a gestire questa mole fenomenale di informazioni? Come riescono ad aiutare gli utenti a trovare quello che cercano? Una panoramica su uno dei più grandi successi del Web.

Il Web introdotto da Tim Berners-Lee intorno al 1990, al quale ci siamo così rapidamente abituati, è costituito da documenti ipermediali. Le informazioni sono in linguaggio naturale (vedi riquadro *I linguaggi Web*), anziché in linguaggio informatico, e i testi sono liberamente strutturati con etichette HTML, ad esempio, per i titoli o gli elenchi. I link sui quali si può cliccare, consentono all'utente di scoprire immagini, musica, film. Essi consentono, inoltre, di muoversi da una pagina all'altra liberamente. E per trovare informazioni in questa confusione, che cosa potrebbe esserci di più semplice? Basta usare quella meraviglia tecnologica che è il motore di ricerca. L'utente sceglie alcune parole chiave. Il motore è in grado di selezionare istantaneamente le pagine che contengono queste parole. La magia è che il motore sa anche offrire, tra decine o addirittura migliaia di milioni di possibili pagine, una scelta di poche pagine che spesso contengono proprio quello che l'utente stava cercando.

Le lingue del Web

Nel 2012, l'inglese rappresenta approssimativamente il 45% delle pagine Web, il 7% il tedesco, il francese è la terza lingua con il 5%; il 90% delle lingue del mondo non sono rappresentate su Internet.

Un indice del Web

Tutto inizia con un indice. Si tratta di una lista ordinata come per l'indice alla fine di un libro, che associa a ogni parola un elenco di pagine Web che contengono quella parola. Ad esempio, una voce in questo indice potrebbe essere:

Casablanca → http://www.imdb.com/title/tt0034583/

che indica che la parola "Casablanca" è presente in quella specifica pagina Web del sito imdb, un database online sul cinema. Effettuando una ricerca nell'indice con più parole come "Casablanca Bogart Bergman", verrà restituito un elenco di pagine Web che contengono una qualsiasi di queste tre parole, compresa la pagina precedente. La difficoltà principale sta nella dimensione di questo indice, che è nell'ordine di un terabyte (un trilione di byte = un milione di megabyte) per gestire qualche miliardo di pagine Web. Il server che gestisce un tale indice deve affrontare problemi di dimensioni crescenti:

1. Per indicizzare un numero crescente di pagine, il server necessita di sempre più spazio per memorizzare l'indice e ogni richiesta richiede un costo sempre più alto per essere eseguita.

2. Con l'aumentare del numero di utenti, il server riceve un numero crescente di richieste.

In entrambi i casi, il server viene rapidamente sovraccaricato. Per risolvere questo problema si usa una tecnica fondamentale di informatica, detta di hashing. Le dimensioni del Web richiedono che questa tecnica venga parallelizzata.

Per distribuire l'indice su, diciamo, $K = 10$ cacolatori, $M_1, \ldots, M_{10}$ useremo una funzione H aleatoria che trasforma una parola in un intero compreso tra 1 e 10. (Una tale funzione è detta funzione di hashing.) La macchina numero $H(w)$ sarà quella responsabile della parola

w Ad esempio, i dati relativi alla parola "Francia" verranno memorizzati sulla macchina H ("*Francia*"), diciamo M_7. Se H è perfettamente aleatoria, i dati risulteranno suddivisi con uguale probabilità tra le dieci macchine, il che risolve il primo problema. Supponiamo che qualcuno richieda i dati per la Francia, basterà interrogare soltanto la macchina M_7. Le richieste in questo modo verranno suddivise con uguale probabilità tra le dieci macchine, il che risolve il secondo problema.

La dimensione dei dati o il numero di utenti possono crescere: sarà sufficiente aumentare il numero di macchine per adattare il sistema. Il parallelismo ci permette di affrontare il carico richiesto dalle dimensioni del Web. Ad esempio, Google utilizza migliaia di macchine organizzate in "server farm" e distribuisce le sue decine di server farm ai quattro angoli del mondo.

Come mai tutto ciò funziona? Grazie al parallelismo. È possibile, in generale, prendere un algoritmo qualsiasi e renderlo più veloce quanto si vuole utilizzando più macchine? La risposta è no! La ricerca ha dimostrato che non tutti i problemi sono parallelizzabili in questo modo, o comunque non sono parallelizzabili così semplicemente. Tuttavia abbiamo visto che la gestione dell'indice è un problema molto semplice e facilmente parallelizzabile: possiamo senza dubbio considerare di indicizzare un numero sempre crescente di pagine, decine di miliardi o più.

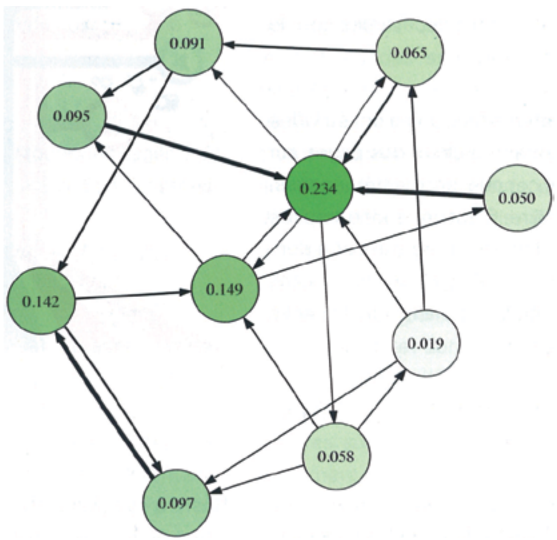


Figura 1. Pagerank: calcolo della popolarità delle immagini sul Web. Ogni cerchio rappresenta un'immagine, le frecce e il loro spessore indicano il numero di collegamenti da una pagina all'altra. I numeri e il colore in ogni cerchio indicano la popolarità.

PageRank

Il cuore del problema rimane ora di scegliere, tra i milioni di pagine che contengono le parole di una interrogazione (*query* in inglese), quelle che sono le più interessanti. Ciò è essenziale, perché un utente raramente si spinge a considerare più dei primi dieci o venti risultati della ricerca.

Inizialmente, motori di ricerca come AltaVista utilizzavano, per classificare le pagine, unicamente delle tecniche statistiche di ricerca dell'informazione. Una pagina veniva considerata più interessante se l'occorrenza di un termine cercato appariva in un titolo o era visualizzata in grassetto. Più un termine era ripetuto nel documento e più alto era il suo "peso". Inoltre si attribuiva un peso maggiore ad un termine raro. Tali tecniche, adeguate su di un piccolo corpus, si sono rivelate deludenti per il Web.

I giovani creatori di Google hanno avuto l'idea di basare la scelta delle pagine su di una tecnica matematica nota, la passeggiata aleatoria. Questa idea, ispirata dai lavori di Jon Kleinberg, è alla base del successo di Google, uno dei successi industriali più incredibili nella storia dell'umanità.

La passeggiata aleatoria

Immaginate un utente ideale del Web con il seguente comportamento. Egli inizia la sua navigazione da una pagina Web a sua scelta, ad esempio, www.inria.fr. Da quella si muove di pagina in pagina nel Web, scegliendo ad ogni passo di cliccare su di uno dei link della pagina corrente. Se la pagina non ha nessun collegamento in uscita, selezionerà un'altra pagina qualsiasi sul Web non importa dove. E, in questo, modo continuerà per un tempo infinito. Qual è la probabilità che egli si trovi su di una data pagina? Questo è ciò che noi definiamo come la popolarità della pagina (vedi figura 1). A priori, una definizione astratta? Un bel concetto matematico del tutto inutile? Non proprio! Nella pratica questa popolarità corrisponde abbastanza bene alla "popolarità" della pagina e a ciò che gli utenti normalmente si aspettano.

Il punto fisso

Resta il problema di calcolare questa popolarità. A questo scopo, bisogna scrivere un'equazione per la quale la popolarità sarà una sua soluzione. Un'equazione, quindi, che codifica il comportamento dell'utente aleatorio. In accordo con il comportamento dell'utente, ogni pagina trasferirà la sua popolarità alle pagine collegate (se una pagina non porta a nessun'altra, essa trasferirà la sua popolarità in misura uguale a tutte le altre pagine). Ignorando alcuni dettagli, questo ci porta a considerare una matrice Q che cattura questi "scambi" di popolarità. Il vettore delle popolarità pop (la lista di tutte le popolarità $pop[i]$ per ogni pagina i) risulta essere la soluzione della seguente equazione di punto fisso:

$$pop = Q \times pop,$$

una notazione molto compatta per un sistema di dieci miliardi di equazioni con dieci miliardi di incognite (una incognita per ogni pagina Web) ...

Abbiamo vinto! Una tecnica già nota ci permetterà di calcolarne la soluzione. Cominciamo dal vettore pop_0 definito da

$$pop_0[i] = 1/N,$$

vale a dire che inizialmente assumiamo che tutte le pagine abbiamo la stessa popolarità. Definiamo:

$$pop_1 = Q \times pop_0; \quad pop_2 = Q \times pop_1;$$
$$pop_3 = Q \times pop_2; \ \ldots$$

Continuando in questo modo, si arriva piuttosto rapidamente ad un punto fisso che risulta essere la soluzione della nostra equazione. Abbiamo quindi calcolato il vettore delle popolarità! (In pratica sei o sette iterazioni sono sufficienti per raggiungere una convergenza soddisfacente.)

Effettuare questo calcolo in modo efficace, con tali volumi di dati, richiede di ingegnerizzare algoritmi molto sofisticati. Forse non dobbiamo considerarla matematica, ma è comunque informatica di grande bellezza. Io stesso ho eseguito una scansione del Web e ho implementato l'algoritmo di PageRank insieme ai miei studenti. Questa è stata una delle mie esperienze più stimolanti come ricercatore.

Altri problemi di ricerca

Abbiamo presentato una versione molto semplificata di un motore di ricerca. Per esempio, quello di Google considera il testo vicino ad un link che punta verso un'altra pagina come parte del contenuto informativo di questa pagina. Questo è quello che viene sfruttato dal "bombing" (vedi il riquadro *Trovare Chuck Norris*). I motori di ricerca sono diventati sempre più sofisticati nel tempo, non fosse altro che per contrastare gli attacchi di "bombers" che cercano di manipolare il Web o degli "spammer" che barano per essere più visibili. L'attuale algoritmo di Google utilizza decine di criteri tra loro combinati in una formula segretamente custodita. Nonostante questo, i motori di ricerca presentano ancora gravi problemi. Ne citiamo alcuni:

- Una misura come il PageRank favorisce la popolarità e quindi ha l'effetto di incoraggiare l'uniformità: le pagine popolari diventando sempre più popolari e le altre sprofondano nell'anonimato. È davvero auspicabile?

- Per interrogare il motore di ricerca, si usa una serie di parole chiave, un "linguaggio" primitivo praticamente senza grammatica. Non si potrebbe fare di meglio?

- È ammissibile escludere certe pagine perché hanno contenuti razzisti, volgari, falso, o per favorire un cliente, o per non dispiacere ad un governo?

- C'è qualcosa di estremamente imbarazzante nel considerevole potere dei motori di ricerca ottenuto attraverso il controllo delle informazioni. Dobbiamo fidarci di loro senza capire i loro meccanismi di classificazione? E perché mantenere questo segreto?

Per gestire i volumi d'informazione del Web, vengono proposte nuove tecniche che fanno un uso sempre maggiore del parallelismo. Numerose tecniche sofisticate vengono progettate per valutare, classificare, filtrare le informazioni. Potremmo citare, ad esempio, le tecniche che sfruttano i sistemi di valutazione (l'utente è invitato a valutare dei servizi come avviene in eBay), i sistemi di raccomandazione automatici (come in Netflix o Meetic), i sistemi che valutano le competenze degli utenti di Internet (come in Mechanikal Turk). Quest'ambito è straordinariamente attivo.

Trovare Chuck Norris

Al momento della stesura di questo articolo, se si digita "trovare Chuck Norris" sul motore di ricerca Google, il primo risultato porta a una pagina che dice " Google non cercherà Chuck Norris perché sa che nessuno può trovare Chuck Norris, sarà lui a trovare voi."

Per ottenere questo risultato, molti utenti di Internet hanno creato pagine Web che puntano a questa pagina con il titolo "trovare Chuck Norris". Arrivando in questo modo a ridefinire il significato della frase "trovare Chuck Norris".

Solo un punto fisso e qualche algoritmo

Mi trovavo nel gruppo di ricerca sui sistemi di informazione a Stanford nel 1995, quando due giovani studenti, Sergei Brin e Larry Page stavano lavorando sul loro prototipo del motore di ricerca Google. Avevano sviluppato gli algoritmi necessari, ma la loro proposta sarebbe potuta passare per stravagante. Infatti, il loro metodo sarebbe stato irrealistico appena qualche anno prima, quando le dimensioni dei supporti di memoria avrebbero richiesto un numero inverosimile di macchine molto costose. Nella metà degli anni '90 ciò è diventato possibile con un numero ragionevole di macchine economiche. In seguito, gli utenti hanno reso popolare il loro motore di ricerca. Come base per questo straordinario successo, si potrebbe parlare dell'eccezionale impresa ingegneristica di far funzionare migliaia di macchine ventiquattro ore su ventiquattro, di modelli di business rivoluzionari, di tecniche di management innovative basate sul culto della creatività. Per quello che mi riguarda, preferisco ricordare che all'inizio si trattava soltanto di un punto fisso e di qualche algoritmo.

Per saperne di più

Brin S., Page L., (1998). *The Anatomy of a Large-Scale Hypertextual Web Search Engine*. WWW Conference.

Abiteboul S., Manolescu I., Rigaux P., Rousset M.-C., Senellart P., *Web data management*, Cambridge University Press. http://webdam.inria.fr/Jorge

Kleinberg J., (1999). *Authoritative sources in a hyperlinked environment*. Journal of the ACM.

Abiteboul S., Preda M., Cobena G., (2003). *Adaptive On-Line Page Importance Computation*. WWW Conference.

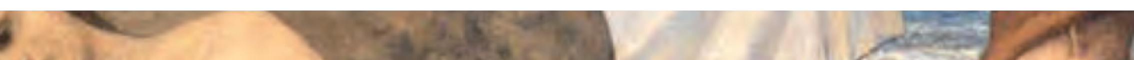
La carriola di Monge e il trasporto ottimale

Yann Brenier, *direttore di ricerca CNRS presso École Polytechnique*

Nato da problemi concreti - come quello di spostare un mucchio di sabbia in maniera efficiente - il trasporto ottimale è uno strumento che ha applicazioni sia all'interno della matematica (dalla geometria all'analisi funzionale) che in altri settori, come, ad esempio, la gestione delle risorse.

Nel XVIII secolo, il matematico francese Gaspard Monge, nel suo trattato *Mémoire sur la théorie des déblais et des remblais* ("Memoria sulla teoria degli scavi e dei detriti"), studiò un problema tra i più concreti (spostare nel modo più efficiente un mucchio di sabbia!) applicando un metodo rigoroso, "ottimale", diremmo oggi. Questo problema è uno dei primi esempi di *ricerca operativa*, il ramo della matematica che si occupa di metodi per la soluzione di problemi combinatorici in modo efficiente (vedi riquadro *La combinatoria*), che è ancora oggi una teoria molto vivace.

In che modo il problema (che potremmo chiamare della "carriola di Monge", avente come obiettivo quello di spostare la sabbia) è un problema combinatorico? Lo possiamo spiegare nel seguente modo. Abbiamo un mucchio di sabbia, composto da piccoli elementi (un elemento può essere un granello o una manciata). Abbiamo, inoltre, una buca da riempire, della quale conosciamo la forma e il cui volume assumiamo essere uguale a quello della sabbia. Possiamo pensare la buca come composta da piccole scatole, ciascuna delle quali è capace di contenere precisamente un elemento di sabbia. La domanda è: dove collocare ogni elemento per ridurre al minimo il percorso totale? Possiamo cercare di rispondere partendo da una descrizione combinatoria del problema, ma, in alcuni casi, può essere più utile dare una descrizione diversa, in base alla forma del profilo del mucchio e della buca, consi-



derate come funzioni. Come vedremo, in questo modo il nostro problema di ottimizzazione potrà essere impiegato per altre applicazioni, non solo per lo spostamento della sabbia.

Altre applicazioni

Il problema del trasporto è stato rivisitato nel 1940 da Leonid Kantorovich (uno dei pochi matematici vincitori del Premio Nobel per l'Economia) a proposito di una questione di allocazione ottimale delle risorse. Nel suo lavoro cerca di risolvere un problema di questo tipo: come utilizzare gli strumenti di produzione (lavoro, capitale ...) per raggiungere la massima soddisfazione dei bisogni? Supponendo, ad esempio, di avere delle miniere di metallo e delle fabbriche con operai, possiamo cercare di determinare come le miniere devono approvvigionare gli impianti delle fabbriche per ridurre al minimo le spese dovute al trasporto del materiale: questo è, in definitiva, lo stesso problema del mucchio di sabbia.

Molte altre questioni (code d'attesa, gestione delle scorte, ecc.), che non affronteremo in dettaglio, appartengono a questo dominio di problemi. Si tratta sempre di trovare come passare da una situazione iniziale ad una situazione finale desiderata nel miglior modo possibile, cioè trovando il percorso migliore nel complesso spazio di tutte le scelte possibili.

Spostare una nuvola di punti

Il punto chiave della teoria è probabilmente il concetto di *geodetica*, parola usata per indicare il percorso più breve tra due punti. Consideriamo due punti nel piano: qual è la strada più breve che li collega? La risposta è ben nota, il segmento di retta. E in una sfera? In questo caso sono gli archi di cerchio massimo: equatore, meridiani (ma non i paralleli); lo si vede, del resto, guardando alle traiettorie dei voli transoceanici.

Ciò che vale per alcuni punti può essere vero per una "nuvola di punti" o per altri oggetti: manipolarli al meglio si riduce a trovare il percorso migliore per spostare tutti questi oggetti insieme da uno stato iniziale ad uno stato finale desiderato.

Per cominciare, consideriamo un esempio: il caso di una nuvola finita formata da quattro punti nel piano (vedi figura 1).

Supponiamo data la configurazione dei punti nello stato iniziale e in quello finale e vediamo come realizzare in modo ottimale il passaggio tra le due configurazioni. Al momento iniziale, la nube è definita da un elenco di quattro punti A_1, A_2, A_3, A_4, e in quello finale altre quattro punti B_1, B_2, B_3, B_4.

Un punto chiave che è necessario capire, è che non teniamo traccia di queste "particelle" singolarmente e che la loro numerazione è arbitraria. Ad esempio, è possibile trasportare, A_1 su B_1, A_2 su B_3, A_3 su B_4 e A_4 su B_2. (Tuttavia, non è consentito, ad esempio, muovere A_1 e A_2 su B_1 lasciando B_4 vacante.) Bisognerà, quindi, cercare una soluzione ottimale tra le possibili permutazioni tra i punti.

Una permutazione è un modo per associare ad ogni punto A un punto B, in modo che ad ogni punto di partenza sia collegato uno ed un solo punto di arrivo. Il numero delle possibili permutazioni dei primi N numeri interi è chiamato fattoriale di N ed è scritto $N!$ Per $N = 4$ il fattoriale vale 24 (e vale più di 20 miliardi per $N = 15$).

La permutazione σ corrispondente al nostro esempio (figura al centro) è σ(1) = 1, σ(2) = 3, σ(3) = 4, σ(4) = 2, mentre la combinazione ottimale è σ(1) = 4, σ(2) = 2, σ(3) = 3, σ(4) = 1.

Trovare la soluzione ottimale è la stessa cosa che minimizzare un *costo* che viene valutato, caso per caso, a seconda delle applicazioni considerate. Il costo di Monge è dato dalla somma di tutte le distanze tra i punti di partenza e i corrispondenti punti d'arrivo. Sono possibili altri tipi di costo. Ad esempio, potremmo prendere in considerazione la somma dei quadrati delle distanze (quello che viene chiamato il costo *quadratico*).

Un problema di ottimizzazione combinatoria

Per come lo abbiamo formulato, il nostro problema di trasporto ottimale appartiene ad un ramo importante della matematica: l'ottimizzazione combinatoria. In questa disciplina, il nostro problema è considerato "facile". Esistono infatti algoritmi che trovano la disposizione ottimale con un numero di operazioni proporzionale al cubo del numero di punti (il che significa che con un computer in cui le operazioni vengono eseguite sequenzialmente, senza parallelismo, il tempo calcolo per N punti sarà più o meno una costante moltiplicata per N^3).

Ricordiamo che il numero di permutazioni è dato da $N! = N \times (N-1) \times \ldots \times 3 \times 2 \times 1$, è, quindi, un fatto notevole riuscire a rendere il tempo di calcolo cubico rispetto a N (infatti,

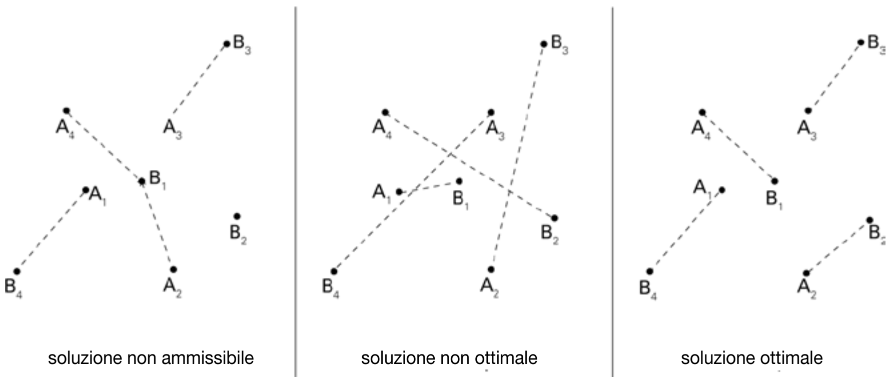


Figura 1. Il trasporto di una nuvola di quattro punti

se per $N = 4$ il fattoriale 24 è ancora un numero modesto, quando N è grande, il fattoriale aumenta quasi come N^N, quindi, in ogni caso, molto più velocemente di N^3, e diventa rapidamente una cifra astronomica).

Tuttavia, anche se è stato osservato che N^3 è molto più piccolo di $N!$, in pratica, gli algoritmi noti sono poco efficienti quando N diventa grande (diciamo oltre le migliaia). Si può sperare di trovare un algoritmo il cui costo computazionale aumenti approssimativamente in maniera proporzionale ad N. Le applicazioni sarebbero spettacolari. Ma per realizzare questo sogno, è necessario un cambiamento di prospettiva, e non considerare più un numero limitato di punti, ma piuttosto molti punti o infiniti punti. Si passa allora da un problema *discreto* a un problema *continuo*.

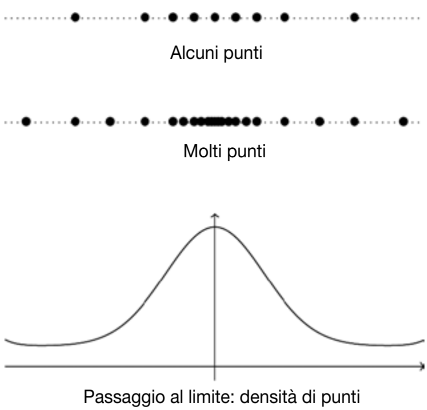


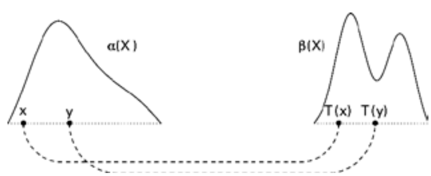
Figura 2. Una nuvola continua di punti

> *Si può sperare di trovare un algoritmo il cui costo computazionale aumenti approssimativamente in proporzione a N. Le applicazioni sarebbero spettacolari.*

Trasporto ottimale di una nuvola continua

Come spesso avviene in matematica e fisica, il limite continuo del problema discreto discusso in precedenza consente di mettere in gioco tutta la forza del calcolo differenziale e integrale (in particolare, le somme che si effettuano nel caso discreto vengono sostituite con degli integrali nel caso continuo).

Da ora in poi, la nuvola di punti sarà descritta da una funzione (densità di probabilità) che, in ogni punto X nello spazio, indica la densità della nuvola. Indichiamo $\alpha(X)$ la densità iniziale e con $\beta(X)$ quella finale. Queste saranno due funzioni a valori reali positivi, il cui integrale su tutto lo spazio è 1 e per le quali l'integrale su ogni sottoinsieme A è la proporzione dei punti della nuvola situati in A. In questa formalizzazione nel continuo il trasporto ottimale sarà una funzione T che invia i punti di partenza in quelli d'arrivo.



Nel caso continuo, se consideriamo un costo quadratico, vale a dire, l'integrale del quadrato della distanza tra il punto sulla nuvola iniziale e quello sulla nuvola finale (invece della somma dei quadrati delle distanze tra i punti di partenza e i corrispondenti punti d'arrivo del caso discreto), certi risultati teorici nella seconda metà degli anni 1980 ci permettono di trovare la soluzione.

Per capire meglio, è utile vedere che cosa accade nel caso di uno spazio ad una dimensione, vale a dire quando la variabile X e le densità α e β si muovono su di una retta. In questo caso, il risultato ci dice che il trasporto ottimale T è il *trasporto monotono crescente*, cioè, quello che invia il punto più a destra della densità α in quello più a destra della densità β e che rispetta l'ordine (se $X \leq Y$ allora $T(X) \leq T(Y)$: si può dimostrare che vi è una sola funzione T con questa proprietà e che rispetta le densità α e β).

In dimensione maggiore di 1, sorge però un problema, che è il seguente: che cosa significa per una funzione di due variabili essere "crescente"? A priori non ha nessun senso. Tuttavia, possiamo usare un trucco: in dimensione 1 la funzione T è crescente se e solo se è la derivata di una funzione convessa (cioè di una funzione il cui grafico è curvato verso l'alto e la cui pendenza aumenta con l'ascissa ed è caratterizzato dal fatto che la derivata seconda è positiva). La nozione di funzione convessa ha senso anche in dimensioni superiori ed è il linguaggio adatto per enunciare il teorema del trasporto ottimale per dimensioni superiori ad uno.

Il risultato degli anni '80 menzionato precedentemente fa intervenire una funzione convessa $\Phi(x,y)$, denominata "potenziale". L'enunciato ci dice che il trasporto ottimale T tra due funzioni di densità α e β è uguale al *gradiente* di questa funzione $\Phi(x,y)$, denotato $D\Phi(x,y)$, che è il vettore le cui componenti sono, in ogni punto, la derivata di Φ rispetto a x e quella rispetto a y.

Resta da trovare la funzione "potenziale" Φ che permette di definire la traiettoria ottimale di ogni particella. Esprimendo il fatto che T invia la densità α sulla densità β, otteniamo delle equazioni differenziali (cioè che coinvolgono delle derivate). In dimensione 1 l'equazione impone un'uguaglianza sulla derivata di T, quindi la derivata seconda di Φ. In dimensioni superiori, si ottiene una formula che esprime il valore del determinante della matrice data dalle derivate seconde di Φ, che deve essere uguale al rapporto tra le densità α e β. Questa equazione ha una soluzione unica Φ e permette quindi di trovare sia il potenziale che la funzione di trasporto T.

Quindi, paradossalmente, il fatto di aver "complicato" il problema (passando dal discreto al continuo) ci ha permesso di risolverlo.

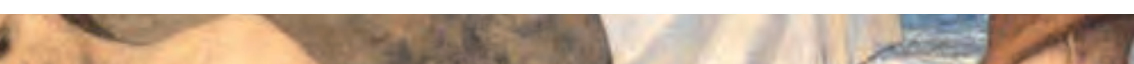
Un'applicazione inaspettata: modificare le immagini

Questo problema essenzialmente matematico ha molte applicazioni pratiche, come mostra il seguente esempio studiato negl'anni 2000 dal team di Jean-Michel Morel presso l'*École Normale Supérieure de Cachan*

Ci viene data un'immagine digitalizzata di un quadro i cui colori sono sbiaditi. Questi colori sono digitalmente codificati in base ai colori elementari blu, rosso e verde. Ad ogni pixel viene assegnata una certa proporzione di blu, di rosso o di verde. Collocando gli N pixel dell'immagine digitale otteniamo una nuvola di N punti nello spazio tridimensionale dei colori. (Questo è uno spazio astratto, non corrispondente al nostro spazio fisico.) Per altre vie, abbiamo un'idea della tavolozza dei colori dell'autore, ottenuta considerando le sue altre opere. Se siamo in grado di fare un'analisi statistica ragionevole, saremo in grado di impostare una nuvola "di riferimento", costituita da N punti nello stesso spazio del colore. Una tecnica di miglioramento del colore consiste nell'effettuare il trasporto ottimale nello spazio colore, dalla nuvola di punti dell'immagine sbiadita alla nuvola di punti di riferimento. Da questo possiamo trasferire di conseguenza i valori dei colori per ciascun pixel.

Si noti che, in questo esempio, è grazie alla formulazione continua che abbiamo potuto studiare il problema, indipendentemente dal fatto che il calcolo numerico venga effettuato su un numero finito di punti!

"Ragazze al mare" P. Puvis de Chavanne



Un’altra applicazione ... all’interno della matematica stessa

Dagli anni ‘90, il trasporto ottimale è divenuto uno strumento di dimostrazione potente ed elegante per dimostrare molte “disuguaglianze” in geometria e analisi funzionale.

Il caso più semplice è quello della famosa “disuguaglianza isoperimetrica”, soluzione del problema di Didone menzionato anche da Virgilio, che può essere enunciato come segue: tra tutte le curve chiuse di lunghezza data, quella che circonda l’area più grande è la circonferenza. Stando alla leggenda, Didone aveva ottenuto dal re locale delle terre, “tante quante ne poteva racchiudere con la pelle di un bue.” Ella prese una pelle e la tagliò in strisce, facendone una lunga corda, per circondare la più grande area possibile, risolvendo così un “problema isoperimetrico” e fondando la città di Cartagine.

Se prendiamo una figura F nel piano e la confrontiamo ad un cerchio C della stessa area, considerando il trasporto ottimale tra le due densità uniformi di F e di C e utilizzando le proprietà di Φ, si può dimostrare che il perimetro di F è necessariamente maggiore di C. Analogamente, nello spazio, il volume massimo racchiuso da una data superficie è dato dalla sfera.

Questi risultati geometrici erano già ben conosciuti prima del trasporto ottimale e questo ci fornisce soltanto una dimostrazione alternativa Tuttavia, con un metodo simile si sono giustificate disuguaglianze non dimostrate in precedenza.

Matematica Pura e Applicata dialogano

Si chiude così un ciclo durato due secoli, nel quale la matematica pura e applicata hanno dialogato e dove l’informatica teorica ha per-

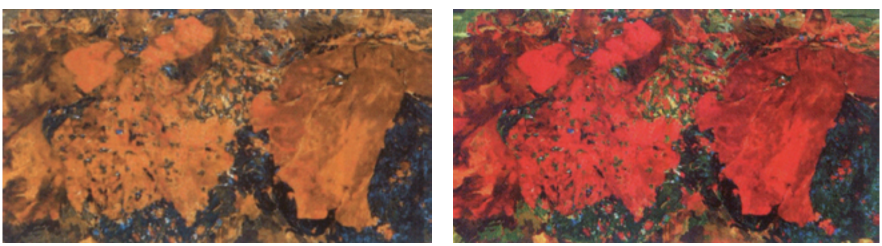


Figura 3: A destra: “Whirlwind” di F. Maliavin. A sinistra: il quadro di F. Maliavin ridipinto con i colori di Puvis de Chavanne, attraverso dei metodi di trasporto ottimale. Fonte Delon-Solomon Sobolevskij, SIAM J. Appl. Math.

messo di capire meglio quello che si può calcolare in modo efficiente. E, soprattutto, dove "cambiando punto di vista" (qui, da un problema discreto a un problema continuo) da un problema complicato si è potuto trovare una soluzione inaspettata.

Per saperne di più

Questo articolo è ispirato quello dello stesso autore pubblicato da *Interstices*.

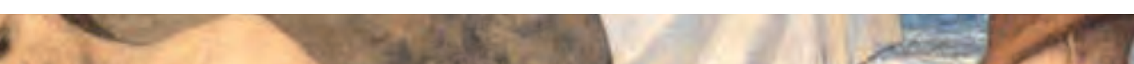
La combinatorica

La combinatorica è il dominio dei problemi della matematica nel quale si elencano e si classificano i possibili modi di far corrispondere o di organizzare un numero finito, ma spesso molto grande, di oggetti. Per esempio:

Quante targhe differenti di tipo LL CCC LL (L lettera, C cifra) si possono emettere?

E se vogliamo che due targhe diverse si distinguano per almeno due caratteri diversi?

Oppure, qual è il numero minimo di colori per colorare una mappa se vogliamo che due stati confinanti abbiano colori diversi?



La statistica per l'individuazione delle alterazioni cromosomiche

Emilie Lebarbier, *ricercatrice*
Stéphane Robin, *direttrice di ricerca INRA presso AgroParisTech*

Le alterazioni cromosomiche sono responsabili di molte malattie, tra cui alcuni tipi di cancro. La rilevazione di piccole alterazioni, essenziale per la diagnosi medica, fa appello ad modello classico in statistica: la segmentazione.

In ogni cellula del corpo umano, ciascun cromosoma è presente in due esemplari o *copie*. Una deviazione da questa regola (*perdita* o *guadagno*) porta ad uno squilibrio, che può causare alcune malattie. La *perdita* corrisponde all'assenza, o alla presenza di una singola copia, di un cromosoma e il *guadagno* sulla presenza di tre, quattro o più copie. Questo squilibrio nel numero di copie è chiamato *alterazione cromosomica*. La più nota è la trisomia, dove un cromosoma è presente in tre copie. Ad esempio, la sindrome di Down (nota anche come *mongolismo*) è causata dalla presenza di tre copie complete del cromosoma 21. È possibile rilevare tali alterazioni per mezzo del *cariotipo* tradizionale.

Altre malattie sono associate ad alterazioni che non influenzano interi cromosomi, ma solo parti di essi. Certe parti vengono "perse" (sono assenti o presenti in una sola copia) o "amplificate" (presente in tre o quattro copie o più). Questo è in particolare la causa di un gran numero di tumori. Le cellule del tessuto malato mostrano tipicamente perdite di regioni che contengono geni "soppressori del tumore" o, al contrario, amplificazione di aree contenenti degli *oncogeni* che favoriscono lo sviluppo del tumore.

Quando sono troppo piccole, queste alterazioni possono purtroppo passare inosservate nello studio del cariotipo. La loro individuazione è pertanto una questione fondamentale nella ricerca medica per identificare i geni coinvolti e per cercare di capire il loro coinvolgimento nello sviluppo delle malattie e, quindi, individuare terapie appropriate.

Il rilevamento del danno cromosomico è una questione critica nella ricerca medica.

Valutare il numero di copie di una porzione del cromosoma

Fu nel 1990 che lo studio di queste piccole modifiche è stato reso possibile con l'avvento delle tecniche di biologia molecolare. Il principio è quello di contare il numero di copie relativo a diversi geni la cui posizione (chiamato *locus*, dal latino) sul genoma è nota, confrontando il risultato tra un individuo malato e uno sano. Questo principio è illustrato con un esempio molto semplice in figura 1a, per cinque *locus*.

Si osservano due alterazioni nel soggetto malato (rosso) rispetto al soggetto sano (verde), che ha precisamente due copie di ogni gene: una porzione persa al *locus* 2 che dà un rapporto tra il numero di copie pari a 1/2 ed un guadagno al *locus* 4 che dà 3/2. Se guardiamo ad un numero più alto di *locus* (figura 1b), di solito molti *locus* successivi possono avere lo stesso status biologico - normale, guadagno o perdita - e quindi formare delle *regioni*. In realtà, non è

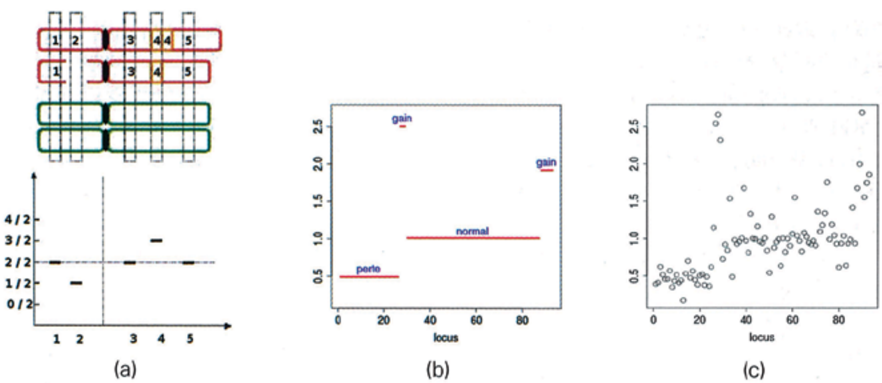
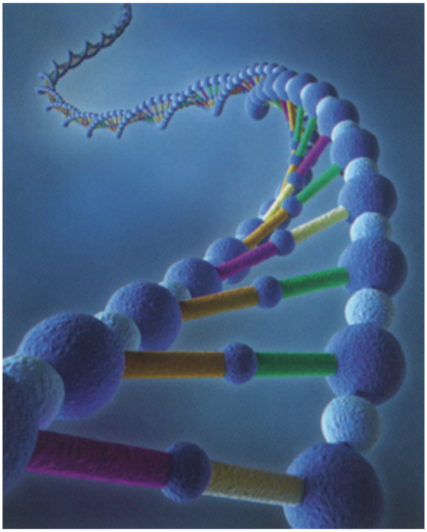


Figura 1. (a) Rapporti "teorici" tra il numero di copie in un individuo malato e quello in un individuo sano (2 copie) per ciascun locus *di cinque* locus *fittizi. (b) I rapporti teorici scomposti in regioni ottenute con l'analisi statistica per 93* locus *reali. (c) Rapporti "empirici" delle copie misurate sperimentalmente per ogni* locus *di questi 93* locus*.*

possibile calcolare direttamente il numero effettivo di copie, ma gli esperimenti biologici ci permettono di ottenere tale numero a meno di un certo errore (errore di misura, variabilità naturale), come mostrato nella figura 1c. Il rapporto tra il numero di copie così ottenuto per differenti *locus* non sarà esattamente uguale all'insieme dei valori teoricamente possibili 1/2, 3/2, 1, 2, ... il che rende impossibile con un controllo manuale.

L'obiettivo statistico è quello di trovare i valori "veri" a partire dai dati osservati. In particolare, come notato sopra, si tratta di individuare e localizzare automaticamente le regioni formate da molti *locus* successivi o da un solo *locus*. Una volta completata l'analisi, il medico può effettuare la sua diagnosi basandosi sulla figura 1b anziché sulla figura 1c.



Un modello matematico

Come spesso accade in statistica, la risoluzione di questo problema passa per la definizione di un modello, cioè, di una traduzione più fedele possibile del processo biologico in linguaggio matematico (sufficientemente semplice da permetterne una risoluzione matematica). Un modello convenzionale statistico specifico per l'individuazione di regioni (o segmenti) è il modello di *segmentazione*. Esso consiste nell'assumere che il valore reale del numero di copie sia lo stesso all'interno di ciascuna regione, che questo valore cambi quando si cambia zona e che il valore osservato in un dato *locus* sia uguale al valore reale della zona a cui appartiene con l'aggiunta di un errore casuale. L'obiettivo statistico è quindi di rispondere a tre domande: Quante regioni ci sono? Dove sono ubicate (vale a dire, quali sono i confini tra le regioni)? Quali sono i valori reali di ogni regione?

Il terzo problema si risolve in modo semplice: viene stimato il vero valore di ogni regione con la media delle osservazioni dirette all'interno della regione. La soluzione statistica delle altre due domande procede in senso inverso. Per un numero fisso di regioni – diciamo k regioni – cerchiamo il posizionamento di queste regioni che meglio si adatta ai dati. Si tratta di un problema algoritmico. La risoluzione fornirà un algoritmo che ci darà la "migliore segmentazione" dei dati in k regioni.

Ovviamente, in generale, il valore di k non è noto. Anche se si ha una informazione a priori sui dati, è difficile fissarlo in anticipo. Dobbiamo quindi determinare, o, più precisamente, scegliere, ancora una volta, il “migliore” k possibile. Se conosciamo la “migliore” segmentazione dei dati per diversi valori di k, potremo allora prenderne uno tra di essi. Questa volta, il problema è in gran parte di natura statistica.

Come ottenere la migliore segmentazione

Sia n il numero di *locus* e stabiliamo che k sia il numero di regioni fissato. Esistono, naturalmente, diverse possibili segmentazioni di n *locus* in k regioni e si tratta di trovare la “migliore”. La figura 2a mostra due esempi di segmentazioni, una in blu e l’altra in rosso, sui dati raccolti lungo il cromosoma 6 di un paziente affetto da un cancro della vescica. Tra queste due segmentazioni, è preferibile la segmentazione blu a quella rossa. Infatti, i dati nella regione “blu” (vale a dire, delimitati dalla segmentazione blu) presentano una minore dispersione intorno al loro “valore vero” rispetto a quelli all’interno della regione “rossa”. La segmentazione blu, quindi, è più aderente ai dati rispetto alla segmentazione rossa. Possiamo quindi scegliere questo criterio di adeguatezza, classico in statistica, come misura della qualità della segmentazione. Resta da provare tutte le possibili segmentazioni e prendere solo le migliori.

Sorge però un problema: quello del numero totale di segmentazioni. È facilmente dimostrabile che, se osserviamo n *locus* cercando di segmentare i dati in k regioni, si hanno tante segmentazioni possibili quante sono le combinazioni di (k-1) oggetti su (n-1), indicate con $C(n\text{-}1,k\text{-}1)$. Ad esempio, per n = 1000 *locus* e k = 10 segmenti, ci sono più di 10^{138} possibili

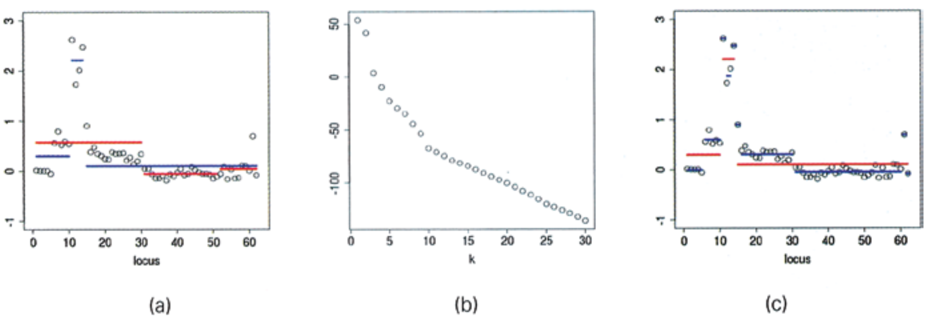


Figura 2. (a) due segmentazioni in tre regioni. (a) Evoluzione dell’adeguatezza in funzione del numero di regioni. (c) Migliori segmentazioni in tre e in dieci regioni.

segmentazioni e neache i più potenti computer possono esplorare tutte queste combinazioni. Una soluzione, tuttavia, è possibile riformulando leggermente il problema.

La migliore segmentazione per un dato numero di regioni è quella che meglio si adatta ai dati.

Definiamo il "costo" di una regione come la dispersione dei dati rispetto alla media di questa regione. Diremo che una regione è "buona" se ha un basso costo. Sapendo che una segmentazione è un insieme di k aree della forma $[1; t_1]$, $[t_1+1; t_2]$ ... $[t_{k-2} + 1; t_{k-1}]$, $[t_{k-1} + 1; n]$, la "migliore" segmentazione corrisponde al percorso che permette:

- di congiungere il punto 1 con punto n,
- con $(k$-1) passi, passando dai punti t_1, t_2, ... ,t_{k-2}, t_{k-1} da determinare,
- con il più basso costo totale.

Tale problema è un problema di ricerca del "percorso più breve" che si può risolvere con un algoritmo che richiede $k \times n^2$ operazioni. Per n = 1000 *locus* e k = 10 segmenti è necessario effettuare 10^7 operazioni, che è molto meno di 10^{138}. In figura 2a, la segmentazione blu è la migliore di tutte le possibili segmentazioni in tre regioni.

Scegliere il numero di regioni

Una prima idea potrebbe essere quella di basarsi sul costo delle segmentazioni sopra definito. Per ciascun numero di regioni possibili k (k = 1, 2 ...), si determina la segmentazione di costo inferiore usando l'algoritmo di percorso più breve e si prende la segmentazione con costo inferiore, che denoteremo $J(k)$. Alla fine si sceglie il numero di regioni k che, tra tutte le possibili, conduce al costo più basso. Possiamo dimostrare che questo ragionamento porta sempre a scegliere il maggior numero possibile di regioni. Infatti, il costo $J(k)$, che riflette l'adattamento della segmentazione ai dati, diminuisce con il numero di regioni: quanto più si aumenta il numero di regioni, tanto più la segmentazione risultante si adatterà ai dati (vedere figura 2c). Quindi, con questo metodo, siamo portati a scegliere la segmentazione nella quale ogni *locus* è una regione a sé (per la quale l'adattamento ai dati è perfetto). Ma questa scelta non ha alcun interesse in quanto equivale a fornire al medico la figura 1c al posto della 1b.

Per ottenere risultati interpretabili, è preferibile scegliere una segmentazione dei dati sufficientemente aderente, ma che non sia troppo complessa, vale a dire, con un numero ragionevole di regioni. La figura 2b mostra l'evoluzione del costo $J(k)$ in base al numero di regioni k ed illustra bene questo obiettivo. Infatti, $J(k)$ diminuisce bruscamente per i primi valori di k, il che significa un netto guadagno dell'adattamento, da un certo k in poi, questa diminuzione è meno marcata (aumentando il valore di k, il costo $J(k)$ migliora solo marginalmente, mentre il modello diventa sempre più complesso). Il valore di k corrispondente a questo rallentamento dovrebbe essere una buona scelta. Quindi l'idea è di rendere più costose le segmentazioni che hanno più regioni. Da un punto di vista matematico, questo significa aggiungere al costo $J(k)$ un termine chiamato *penalità*, che tiene conto della complessità della segmentazione.

Resta da definire cosa si intende per "complessità" di una segmentazione con k segmenti. Per descrivere la segmentazione, è necessario determinare i k valori delle medie, ma anche i confini t_k tra le regioni. Abbiamo visto che questa ultima ricerca rende necessario esplorare un numero troppo alto di segmentazioni. Così, una penalità adeguata dovrebbe tenere conto di queste due quantità, solitamente nella forma

$$a \times k + b \times \log\left[C(n-1,\ k-1)\right].$$

Sarebbe illusorio pensare che il criterio così costruito (e che richiede la scelta delle costanti a, b) sia infallibile, cioè che possa portare alla migliore soluzione in qualsiasi situazione. Poter disporre del massimo delle informazioni è, dunque, importante. In aggiunta alle informazioni in possesso del biologo (che conosce meglio di chiunque i dati), una analisi del comportamento del costo $J(k)$ può essere utilizzata per estrarre diverse soluzioni per k che possono essere biologicamente significative. Ad esempio, la figura 2b mostra che $J(k)$ ha due inflessioni per k = 3 e per k = 10. Sebbene il criterio suggerisca la soluzione k = 3, risulta comunque importante studiare la segmentazione per k = 10, dato che, come vedremo tra poco, consente l'individuazione di un'alterazione addizionale.

La migliore segmentazione consente il rilevamento di una alterazione supplementare.

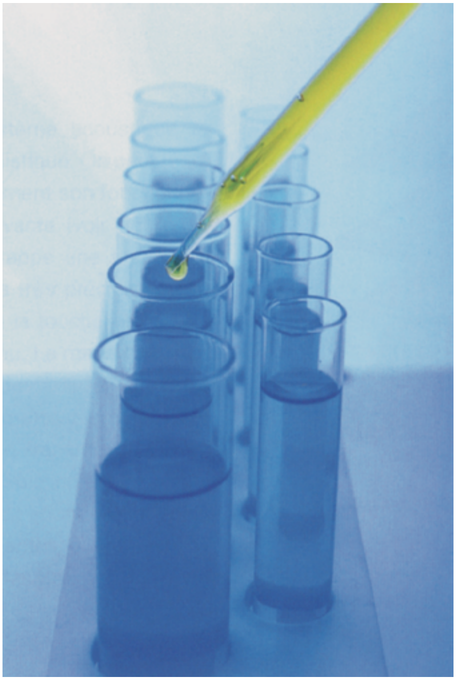
Un rilevamento più fine

I criteri sopra descritti ci portano a considerare k = 10 regioni per i dati presentati in figura 2. La segmentazione associata è quella in blu nella figura 2c. In questo modo viene individuata una regione amplificata (*locus* da 11 a 14). Questa è la regione del cromosoma 6 che contiene il gene di amplificazione E2F3, noto per essere coinvolto nello sviluppo del cancro della vescica. L'amplificazione di questa regione si ritrova anche in altri pazienti con lo stesso tipo di cancro.

Questa segmentazione rileva anche una variazione nel *locus* 61. Questo *locus* è conosciuto dai biologi per essere collegato ad un polimorfismo (una variazione genetica locale specifica del paziente). Non è quindi sorprendente osservare un'alterazione in questo *locus* in alcuni pazienti.

Come accennato sopra, un secondo punto di flesso esiste nel costo $J(k)$ per k = 3. La segmentazione associata nella figura 2c è mostrata in rosso. Questa segmentazione, grossolana rileva già la regione contenente il gene E2F3, ma non il polimorfismo del *locus* 61.

Gli autori ringraziano Theo Robin per la sua attenta lettura e per i suoi commenti.



Il pianoforte che i matematici sognano

Juliette Chabassier, *ricercatrice Inria, gruppo di ricerca Magic-3D*

Come molti fenomeni fisici, il funzionamento di un pianoforte può essere modellizzato matematicamente. Una volta fatto ciò, possiamo andare oltre e sognare pianoforti impossibili o immaginare nuovi suoni. In questo modo la ricerca offre al compositore un meraviglioso campo di esplorazione e di creazione.

Il pianoforte è un sistema acustico e meccanico piuttosto sofisticato. Il suo funzionamento può essere riassunto schematicamente nel modo seguente (vedi figura 1): il dito del pianista batte sul tasto, un meccanismo di precisione propaga il movimento del tasto e mette in moto un martelletto. Il martelletto colpisce da una a tre corde insieme (a seconda della nota scelta), che iniziano a vibrare. Il ponticello permette di trasmettere l'energia dalle corde alla tavola armonica, che entra anch'essa in vibrazione, mettendo in moto le molecole d'aria circostanti e causando la propagazione del suono nell'aria.

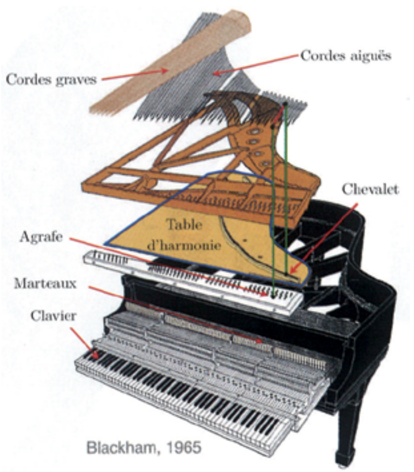


Figura 1. Esploso di un pianoforte a coda.

Le EDP negli accordi

Come la maggior parte dei fenomeni fisici che ci circondano, il funzionamento di un pianoforte può essere descritto matematicamente mediante equazioni differenziali alle derivate parziali (abbreviato EDP o spesso anche PDE dall'acronimo inglese). Le EDP sono delle equazioni che coinvolgono le derivate di una funzione o più funzioni in relazione con le loro derivate rispetto a certe variabili: il tempo, lo spazio ... Alcune EDP sono piuttosto famose: l'*equazione del calore*, che modellizza la dinamica della temperatura in un mezzo fisico; le *equazioni di propagazione delle onde*, che modellizzano la vibrazione di una corda e la propagazione del suono; le *equazioni di Navier-Stokes* che intervengono nella meccanica dei fluidi; le *equazioni di Maxwell*, che sono quelle dell'elettromagnetismo; le *equazioni dell'elastodinamica*, che modellizzano le onde sismiche o la deformazione dei materiali; l'*equazione di Schrödinger*, che modellizza l'evoluzione di una particella quantistica.

Torniamo al nostro pianoforte: ciascuna delle componenti acustiche sopracitate può essere modellizzata da una EDP (per le corde sarà una equazione d'onda unidimensionale, per la tavola armonica, un'equazione di vibrazione della membrana, per l'aria un'equazione di propagazione d'onda tridimensionale) o da un'equazione di natura diversa (per il martelletto ci sarà un'equazione differenziale ordinaria, per il ponticello avremo un'uguaglianza di velocità, e anche per la coppia, tavola armonica/aria, l'equazione esprimerà l'uguaglianza di velocità tra mezzo meccanico e acustico).

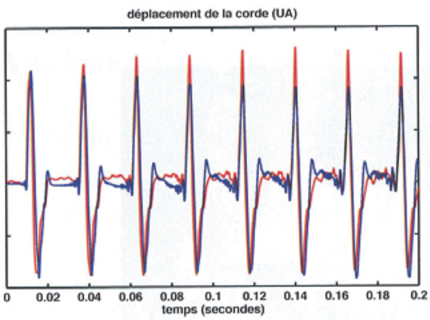
Figura 2. Magliatura di una tavola armonica.

Tipicamente, le EDP sono equazioni complesse per le quali è generalmente impossibile esprimere una soluzione in formula chiusa. I problemi fisici che esse traducono e che spesso combinano molte di queste equazioni, sono necessariamente ancora più complessi. Vari metodi vengono utilizzati per calcolare le soluzioni approssimate che meglio descrivono la realtà (vedi riquadro *Analisi Numerica ed EDP*). Questo, spesso, richiede di discretizzare il problema e far calcolare dal computer le soluzioni approssimate.

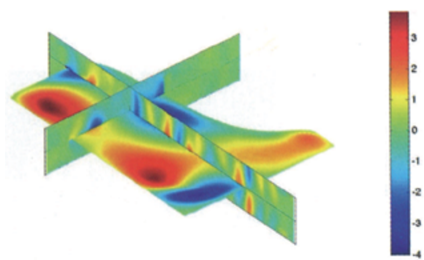
Un modello matematico del pianoforte viene descritto mediante le EDP e successivamente discretizzato attraverso una rete. Infine le soluzioni vengono calcolate numericamente.

Questo è quanto fatto nel caso del pianoforte: un modello matematico viene descritto attraverso le equazioni e successivamente discretizzato attraverso un *magliatura* (si veda la magliatura di una tavola armonica nella figura 2); infine le soluzioni sono calcolate numericamente attraverso metodi di analisi numerica appositamente sviluppati.

Con questo modello e le simulazioni numeriche che ne seguono, si riescono a riprodurre fedelmente le forme d'onda (vedi figura 3), ma, anche, per esempio, conoscere il campo di pressione dell'aria su di un insieme di punti intorno al pianoforte, che sarebbe impossibile misurare senza perturbare il sistema: se si tentasse di fare una tale misurazione si dovrebbe circondare il pianoforte con una rete di sensori la cui presenza potrebbe



(a) confronto tra simulazione numerica (blu) / e dati sperimentali (rosso): moto della corda D#1 durante i primi 200 millisecondi.



(b) Moto della tavola armonica e pressione dell'aria (due sezioni) in Pa (scala dei colori a destra) 16 millisecondi dopo il colpo del martelletto sulla corda C2.

Figura 3. I risultati della simulazione numerica: verifica e previsione.

causare un fenomeno di diffrazione del suono e falsare di conseguenza la misurazione.

Pianoforti virtuali

Una delle più interessanti applicazioni di questo modello e della sua discretizzazione numerica è il contributo alla costruzione degli strumenti. Ancora oggi, la costruzione di un pianoforte si basa soprattutto su di un sapere empirico nato da secoli di sperimentazioni, fallimenti, successi ... Coloro che progettano e costruiscono gli strumenti, hanno acquisito una serie conoscenze estremamente precise. Si trovano, ad esempio, nella letteratura affermazioni come: “I rullini dei martelletti influiscono sulla sensazione del tocco del pianoforte in modo piuttosto decisivo” oppure “la tavola armonica è in abete di Sitka, un legno con densità ed elasticità ideali per la risonanza”. Queste affermazioni incuriosiscono molto i ricercatori in acustica musicale, i quali cercano, con metodi scientifici, di dar loro ragione o torto e di avanzare nella comprensione dei fenomeni coinvolti. Il modello del pianoforte può, quindi, essere utilizzato per isolare alcuni fenomeni, al fine di comprenderne la loro influenza sul suono, sull’irraggiamento, o la trasmissione di energia ... ma anche per costruire virtualmente dei pianoforti che non esistono (modificarne la forma o le dimensioni della tavola, i materiali utilizzati, il loro impiego ...) e ascoltare il suono che avrebbero se fossero costruiti.

Ancora più affascinate è il fatto che il modello è in grado di generare suoni di oggetti che non possono esistere per motivi pratici (costruiti con materiali inventati, corde di sette metri di

lunghezza, pianoforti galleggianti e senza cornice né piedini ...), ma comunque rispettando le leggi della fisica e il cui suono sembra quindi plausibile per l'orecchio (nel senso che il cervello è in grado di associarlo ad un evento, ad una causa fisica, ad una dinamica). Si giunge in questo modo ad una applicazione interessante dal punto di vista dell'interazione tra scienza e musica: la ricerca offre al compositore dei nuovi strumenti, un materiale sonoro finora impossibile e adattabile a piacere. In un'epoca in cui il timbro e la manipolazione del suono sono al centro della creazione musicale, questo gli permette di andare oltre.

Il modello può generare suoni di oggetti che non possono esistere per ragioni pratiche.

Uno strumento per il compositore

Queste opportunità sono ben illustrate dal lavoro del compositore Guillaume Loizillon, che utilizza un software di sintesi sonora basato su modelli fisici sviluppati dall'Istituto per la Ricerca e Coordinamento di Acustica / Musica, Modalys, che consente di modellare una moltitudine di strumenti inesistenti e chimerici. Egli ha detto in proposito: "Quello che trovo interessante in Modalys, più che l'imitazione di un timbro, è che permette, al livello di un unico suono, di dare, in modo molto evidente, l'idea del materiale, del metallo, del legno. Inoltre, aiuta a creare delle dinamiche, degli eventi, non semplicemente degli oggetti. [...] Vale a dire che il suono indica un evento: ecco che rotola, rimbalza."

Analisi Numerica per le EDP

Quando ci concentriamo su delle applicazioni concrete delle EDP (in meteorologia, in sismologia, in aeronautica, nella costruzione di automobili ...), non è sufficiente conoscere l'esistenza e il comportamento di una soluzione in uno spazio matematico astratto. I fisici hanno bisogno almeno di stimare queste soluzioni quantitativamente. La scienza che consiste nel cercare una approssimazione della soluzione di una EDP in modo preciso ed efficiente tramite l'uso del computer, è quella del calcolo scientifico e dell'analisi numerica. L'idea è quella di riportare lo spazio matematico infinito delle soluzioni delle EDP al carattere fondamentalmente finito e discreto (al contrario di continuo) del computer, assicurando, però, che maggiori saranno gli sforzi (in termini di tempo di calcolo o di potenza del processore), più ci avvicineremo alla soluzione "reale". Uno dei metodi più noti e più comuni nel settore industriale è probabilmente il *metodo degli elementi finiti*.

Il compositore può anche oltrepassare il mondo fisico esistente e immaginarne un altro, con leggi diverse "Inoltre, quello che amo è provare modelli più utopici", spiega Guillaume Loizillon. "Con Jean Claude Risset [ndr: anche lui compositore], abbiamo discusso della questione *se siamo in grado di creare una fisica non terrestre.* [...] Possiamo immaginare una fisica in cui gli oggetti rimbalzano per ore e ore, o dove la gravità non è la stessa."

In questo modo, la modellizzazione strumentale apre ai musicisti dei campi di esplorazione incredibili.

Per saperne di più

Sito web: http://modelisation.piano.free.fr

Le citazioni de Guillaume Loizillon provengono da *La synthèse par modèles physiques*, tesi di laurea di Aline Hufschmitt, Université Paris Sorbonne (Paris IV), disponibile all'indirizzo alinehuf3.free.fr

Come far cooperare individui egoisti?

Yannick Viossat, *professore presso l'Università Paris-Dauphine*

La cooperazione è al centro di molti comportamenti sociali e biologici. La teoria dei giochi può spiegare la scelta delle strategie di cooperazione e comprenderne i meccanismi in contesti in cui gli individui sono in concorrenza, dalle situazioni di guerra alla regolamentazione della pesca.

Il mondo vivente offre molti esempi di cooperazione: gli insetti sociali, i versi degli animali per indicare la presenza di un predatore, gli impala che si puliscono a vicenda ... Darwin considerava questi comportamenti cooperativi come un enigma e una sfida alla sua teoria. La selezione naturale non dovrebbe incoraggiare un comportamento egoista? Questo paradosso della cooperazione è al centro dell'analisi delle principali transizioni nell'evoluzione, come la comparsa della vita o quella degli organismi multicellulari. In alcune di queste transizioni, piccole entità cooperano per formare un'entità più complessa. Queste costruzioni sussistono, anche se sembrano esposte all'apparizione di individui parassiti, che sfruttano la cooperazione degli altri, ma che non cooperano essi stessi.

Capire l'origine e il funzionamento della cooperazione nel mondo vivente è un problema centrale in biologia evolutiva.

Interesse collettivo, interessi individuali

Questioni simili sorgono in scienze sociali. Molti conflitti esistono tra l'interesse collettivo e gli interessi individuali. Può, dunque, la cooperazione emergere spontaneamente? Più concretamente: perché continuiamo a pescare troppo e ad esaurire il patrimonio ittico? O viceversa, perché in alcune trincee della prima guerra mondiale, i soldati hanno evitato volutamente di interferire con il rifornimento dei loro nemici?

La matematica permette di fare luce su questi interrogativi, analizzando i modelli di conflitto tra l'interesse collettivo e gli interessi individuali. Il più semplice di questi modelli, chiamato per motivi aneddotici il dilemma del prigioniero, è il seguente: due individui, chiamati giocatore 1 e giocatore 2, devono decidere di aiutarsi a vicenda (cooperare) o di non aiutarsi (tradire). Ogni giocatore deve effettuare la sua scelta senza conoscere la scelta dell'altro.

Cooperare produce un beneficio di 4 all'altro, ma costa 1 a chi coopera. Il tradimento non dà nulla e non costa nulla. Quindi, se i giocatori cooperano tra loro, entrambi guadagnano 4-1 = 3. La tabella seguente riassume le azioni e i corrispondenti guadagni. I guadagni del giocatore 1 (G1) sono in rosso, quelli del giocatore 2 (G2) in blu. Il valore esatto di questi guadagni non è essenziale purché il loro ordine (dal più piccolo al più grande) rimanga invariato.

Molte situazioni della vita quotidiana presentano delle analogie: I bambini che possono prestarsi i loro giocattoli (C) o non prestarseli (T); le grandi potenze durante la Guerra Fredda, potevano limitare i loro armamenti (C) o armarsi sempre di più (T); i cittadini possono riciclare i rifiuti (C) o oppure no (T), ecc.

G2 / G1	Se G2 coopera (C)	Se G2 tradisce (T)
Se G1 coopera (C)	3,3	−1,4
Se G1 tradisce (T)	4, −1	0,0

L'analisi di queste situazioni è quella della teoria dei giochi. Questo ramo della matematica esamina l'interazione tra attori che si comportano secondo una strategia. Riguarda tutte quelle situazioni in cui degli attori devono prendere delle decisioni che ne determinano l'esito.

Assicurare il futuro

La teoria dei giochi ha identificato diversi meccanismi per superare i conflitti tra l'interesse pubblico e gli interessi individuali. Vediamo ora il funzionamento e i limiti di uno di questi meccanismi: la reciprocità diretta (aiutare coloro che ci aiutano).

Altri meccanismi di cooperazione

La reciprocità diretta non è l'unico meccanismo sul quale può basarsi la cooperazione tra individui. Tra gli altri possiamo citare: la creazione di istituzioni incaricate di far rispettare i contratti, la reciprocità indiretta (aiutare coloro che hanno aiutato gli altri) o la selezione per parentela (aiutare le persone che sono legate a noi), fondamentale in biologia.

Ritorniamo al dilemma del prigioniero. Quando il gioco è fatto da una sola partita, la giocata T frutta sempre una unità in più della giocata C (4 invece di 3 se l'altro gioca C, 0 invece di –1 se l'altro gioca T). I due individui sono quindi singolarmente interessati a tradire. Il paradosso sta nel fatto che in questo modo entrambi ottengono 0, mentre se avessero cooperato avrebbero ottenuto 3.

Se invece si fanno diverse partite la situazione cambia: le mie scelte di oggi influenzano le scelte del mio compagno in futuro. Cooperare può essere una forma di investimento: qualcosa che mi fa guadagnare meno nell'immediato può fruttare maggiori guadagni in futuro, se questo mio comportamento renderà l'altro più cooperativo. Tuttavia, se l'altro giocatore tradisce sistematicamente, anch'io ho interesse a tradire. Dunque non esiste una strategia che sia sempre la migliore.

Le mie scelte di oggi influenzano le scelte del mio compagno in futuro.

Tuttavia, alcune strategie producono spesso buoni risultati. La strategia *Do ut des,* in particolare, ha vinto diversi tornei tra computer, dove ogni strategia ha giocato contro le altre e in cui l'obiettivo era quello di avere il più alto punteggio totale. Questa strategia di semplice cooperazione inizia con la cooperazione reciproca e successivamente ripete quello che ha fatto l'altro nel turno precedente.

È non sfruttando gli altri che la *Do ut des* ottiene dei buoni risultati. In realtà, in un duello, non guadagna mai più del suo avversario. Ma riesce a innescare una logica di cooperazione e, in media, ottiene dei punteggi molto buoni. L'analisi dei tornei del dilemma del prigioniero fa emergere alcune caratteristiche delle strategie di successo: cercare di cooperare, non cercare di vincere più degli altri, non lasciarsi sfruttare, perdonare un tradimento episodico, fare scelte comprensibili.

Il sistema del "vivi e lascia vivere" della prima guerra mondiale fornisce un esempio concreto di cooperazione in un contesto difficile. Questo sistema era quello di non attaccare il nemico, purché neanche lui lo facesse. Questo atteggiamento è stato fortemente scoraggiato dal comando militare. Un ufficiale inglese raccontò in questo modo la sua visita alle trincee: “Sono rimasto stupito nel vedere i soldati tedeschi che si muovevano nel raggio d’azione dei nostri fucili. I nostri uomini non sembravano nemmeno accorgersene. (...) Una cosa del genere non dovrebbe essere consentita.” Tuttavia, questa dinamica si instaurò in molte parti del fronte, perché nella guerra di trincea, le stesse unità si affrontavano per periodi abbastanza lunghi per far si che la logica di reciprocità potesse funzionare. Un soldato ha spiegato: “Sarebbe un gioco da ragazzi bombardare le vie di accesso alle trincee, affollata come dovrebbe essere dei mezzi di approvvigionamento (...), ma nel complesso tutto è tranquillo (...). Se si impedisce al nemico di rifornirsi la sua reazione sarà semplice: impedirà il nostro rifornimento.”

Un’analisi precisa del dilemma del prigioniero aiuta a capire meglio i fattori che favoriscono le strategie collaborative. Il guadagno con il tradimento non deve essere troppo alto e dovrà esserci sufficiente tempo a disposizione. Perché questo avvenga, bisogna che l’interazione si ripeta per un periodo sufficientemente lungo. Bisogna anche che i giocatori non sappiano qual è il momento in cui il gioco finisce, altrimenti, all’ultimo turno, i due giocatori hanno convenienza a tradire e, in previsione di ciò, sono portati a tradire al penultimo turno e così via.

La teoria dei giochi evolutivi

Per capire meglio le condizioni che favoriscono l’instaurarsi di strategie di reciprocità, è utile usare la teoria dei giochi evolutivi. Essa studia le situazioni in cui un gran numero di individui interagiscono e fanno valutare il loro comportamento in base al risultato di interazioni precedenti. Si presume che le strategie che ottengono buoni risultati tendano a diffondersi nella popolazione. Una formulazione precisa di questo principio dà luogo ad un sistema dinamico, in grado di modellizzare fenomeni sia di selezione naturale sia di apprendimento.

In un mondo in cui si possono commettere errori, è bene riparare ai danni provocati involontariamente agli altri e non vendicarsi sistematicamente delle aggressioni subite.

Nei modelli di mutazione-selezione, vi è un alternarsi di fasi di tradimento e di cooperazione. Partendo da uno stato in cui nessuno coopera, le mutazioni possono far nascere delle strategie di tipo *Do ut des*. Se queste raggiungono una frequenza sufficientemente elevata, esse ottengono risultati migliori del tradimento e invadono la popolazione. Questo porta a uno stato in cui tutti gli individui giocano con una strategia di cooperazione. La strategia che consiste nel cooperare indiscriminatamente può allora aumentare di frequenza, in quanto non porta ad una contro-selezione. Ma se i cooperatori indiscriminati sono troppo frequenti, la strategia di tradimento è favorita e può diventare predominante fino a quando non comincia un nuovo ciclo.

Ora supponiamo che in ogni fase, con una piccola probabilità, i giocatori non riescano a realizzare l'azione che hanno scelto: essi cooperano invece di tradire, o il contrario. Quando *Do ut des* gioca contro se stesso, il primo errore innesca una faida distruttiva (vedi figura 1). Pertanto, *Do ut des* non può instaurarsi stabilmente nella popolazione

Giocatore 1	C...CTCTC ... TCC
Giocatore 2	C...CCTCT ... CCC

Figura 1. Do ut des contro un altro giocatore Do ut des, con errori (C significa "cooperare" e T "tradire").

Un tradimento involontario (in rosso) porta ad un'alternanza di cooperazione e tradimento. Bisogna aspettare una cooperazione involontaria (in blu) per riprendere la cooperazione. Altre strategie di cooperazione prendono il sopravvento, ad esempio il *Do ut des "pentito"* prevede di compensare un tradimento errato cooperando due volte qualsiasi cosa faccia l'altro oppure il *Do ut des* "generoso" che coopera con probabilità positiva, anche se l'altro ha tradito. Queste piccole differenze possono far superare le crisi causate da un tradimento involontario. In un mondo in cui si possono commettere errori, è bene riparare ai danni provocati involontariamente e non vendicarsi sistematicamente delle aggressioni subite.

La pesca o la tempesta di neve: altri tipologie di modelli

Fin qui abbiamo ipotizzato che l'interazione si ripetesse, ma che il valore delle vincite durante il gioco non mutasse. Ora, immaginate un modello volontario per la regolamentazione della pesca, dove un gran numero di attori deve decidere la quantità di pesce che si può pescare ogni anno. Pescare poco permette alla fauna ittica di rinnovarsi, ed è analogo a cooperare. Pescare il più possibile esaurisce le risorse, ed è analogo al tradimento. Ogni attore ha una piccola influenza sull'evoluzione delle scorte ittiche, il suo interesse individuale è quello di pescare il più possibile, anche se, collettivamente, è consigliabile limitare lo sfruttamento. La situazione è analoga a quella del dilemma del prigioniero.

Tuttavia, una rappresentazione più accurata dovrebbe tener conto dell'evoluzione dell'abbondanza di pesce. "Pesca il più possibile", per esempio, frutta sempre meno via via che le scorte si esauriscono. Il guadagno dei giocatori in un dato momento dipende quindi, non solo dalle loro azioni presenti, ma anche dallo stato generale del sistema, che nel nostro caso è la riserva di pesce. Questo stato si evolve in base alle azioni dei diversi attori e dei possibilil rischi. Lo studio di questo tipo di situazione richiede modelli più avanzati, che conducono ai *giochi stocastici*.

Un'altra ipotesi che abbiamo implicitamente utilizzato è che un individuo abbia uguale possibilità di interagire con ciascuno degli altri individui. In realtà, interagiamo in primo luogo con i nostri colleghi, i nostri amici o nostri vicini. Questo può essere preso in considerazione assumendo che gli individui siano collocati in uno spazio e interagiscano principalmente con i loro vicini. L'analisi di questi modelli ha mostrato che nel dilemma del prigioniero la cooperazione emerge più facilmente quando le interazioni sono locali rispetto ai casi in cui gli individui interagiscono con qualsiasi altro individuo.

Allo stesso modo è istruttivo studiare modelli dove il conflitto tra interessi individuali e interesse pubblico non è così brutale. Nel gioco della tempesta di neve, per esempio, due guidatori trovano la strada interrotta per la neve. Ognuno ha una pala nella sua auto, e può sia sgombrare la strada (cooperare) o aspettare nella sua auto (tradire) sperando che l'altro lo faccia da solo. La cosa preferibile è che l'altro sgombri la strada da solo o in subordine, sgombrarla insieme in due, ed è comunque preferibile sgombrarla da solo, piuttosto che passare la notte nella propria auto, cosa che accadrà se la strada non viene sgombrata. Nonostante ci sia sempre un conflitto tra l'interesse collettivo (sgombrare la strada) e l'interesse individuale (lasciar fare all'altro), la situazione non è più quella del dilemma del prigioniero. Infatti, tradire, non è sempre conveniente: è così solo se gli altri cooperano. L'analisi di questo gioco rivela differenze più sottili. Come abbiamo visto, se gli individui interagiscono maggiormente con i loro vicini, la cooperazione viene favorita nel dilemma del prigioniero. Questo effetto di interazioni locali non è presente nel gioco della tempesta di neve. Più in generale, se il dilemma del prigioniero è stato ampiamente studiato, questo non è avvenuto per altri modelli, ugualmente rilevanti per studiare i fenomeni di cooperazione. Un'analisi più avanzata di questi modelli è necessaria, e in gran parte ancora da svolgere, per capire quali risultati sono validi in generale. Se la matematica aiuta a capire certi fenomeni di cooperazione, ci sono ancora molti gli sforzi da fare per raggiungere una comprensione più completa.

Per saperne di più

Axelrod R., (1992). *Donnant-Donnant, théorie du comportement coopératif*, éditions O. Jacob.

Sigmund K., (2010). *The Calculus of Selfishness*, Princeton University Press.

La pagina web http://www.univie.ac.at/virtuallabs propone programmi liberamente accessibili per l'esplorazione della dinamica dei dilemmi sociali.

Ictus: la matematica viene in soccorso

Emmanuel Grenier, *professore presso École Normale Supérieure de Lyon*

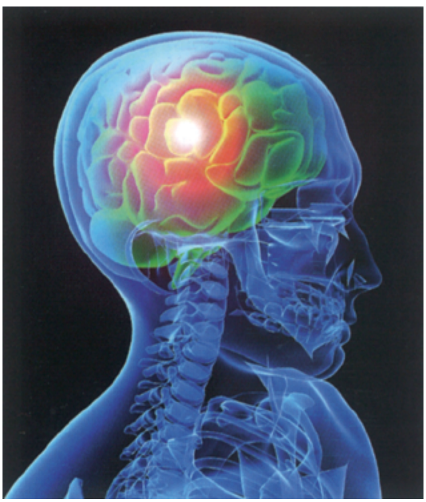
L'ictus, che colpisce migliaia di persone ogni anno, è una malattia complessa la cui diagnosi e il cui trattamento devono ancora essere migliorati. Questo è un settore nel quale la modellizzazione matematica può aiutare la ricerca medica, in particolare affiancandosi alla sperimentazione sugli animali.

Secondo l'associazione di assistenza ai pazienti e alle famiglie colpite da ictus, la Francia conta 150.000 nuovi casi di ictus ogni anno. L'ictus ischemico inizia con l'ostruzione di un'arteria del cervello, ad esempio con un piccolo coagulo di sangue: l'area del cervello normalmente irrigata da tale arteria, viene privata del suo abituale flusso di sangue. Ai neuroni in questa zona viene a mancare ciò che è necessario al loro normale funzionamento e muoiono. Ma non è tutto: la presenza di cellule morte (e in particolare necrotizzate, cioè esplose) provoca disfunzioni organiche ed è, quindi, necessario rimuoverle. Sorge un'infiammazione nell'area interessata, causando essa stessa danneggiamenti, provocando gonfiore dei tessuti e rilasciando sostanze tossiche. Questi danni si vanno ad aggiungere a quelli già causati dall'ictus. Inoltre, i radicali liberi normalmente contenuti nelle cellule, ma che, a seguito dello scoppio di queste, si trovano nello spazio extracellulare, si ossidano, producendo altre sostanze tossiche. I danni possono così susseguirsi per diversi giorni o anche per settimane dopo l'inizio dell'incidente. L'ictus è, dunque, una condizione molto complessa, che coinvolge molteplici attori. Secondo le zone colpite, il paziente soffre di paralisi, afasia, ...

L'arsenale terapeutico a disposizione per curare ictus è purtroppo limitato. I medici cercano di sbloccare l'arteria interessata per ripristinare al più presto un normale flusso di sangue, ma ciò non è esente da rischi. Inoltre, la diagnosi di ictus è di per sé difficile.

Il matematico traduce le conoscenze biologiche e mediche sotto forma di equazioni. Poi, a partire da queste equazioni, effettuerà delle simulazioni al computer.

Per far progredire la ricerca in questo settore, i medici e i biologi lavorano su “modelli animali”, causando ictus nei roditori (ratti, gerbilli) e successivamente osservandoli mediante radiografie o biopsia. Il problema sorge quando si passa alla sperimentazione clinica sull’uomo, in cui i risultati possono essere diversi da quelli sugli animali, e in contrasto con l’obiettivo ricercato, come vedremo più avanti.



Modellizzazione e simulazione dei fenomeni

In che modo la matematica può aiutare la ricerca clinica? Il vantaggio principale della matematica è che permette la quantificazione dei fenomeni. Il medico e il biologo pensano e analizzano i sistemi viventi attraverso certe relazioni, influenze, interazioni ... Ma non possono sempre quantificare con precisione gli effetti, per dar loro un valore numerico.

Di fronte ad un sistema complesso, un biologo o un medico avrà l’intuizione del suo funzionamento. Prevederà degli scenari. Tuttavia, i sistemi viventi sono così complessi che è difficile avere una corretta e completa intuizione del loro funzionamento.

Il matematico, da parte sua, tenta di modellizzare matematicamente i fenomeni. In altre parole, traduce le conoscenze biologiche e mediche sotto forma di equazioni. Poi, da queste equazioni, effettuerà delle simulazioni al computer. Invece di avere un modello animale, si costruirà un modello teorico (equazioni). Invece di esperienze, farà i conti. Certo, il modello astratto e il modello animale sono due approssimazioni della realtà. Il modello astratto non può spiegare la complessità della vita. È intrinsecamente falso, approssimativo, in costante mutamento e tiene conto soltanto di una parte delle conoscenze.

Un matematico non scoprirà una nuova molecola. Ma, esprimendo in equazioni le intuizioni dei medici, ne mette alla prova la coerenza e la validità. A volte i vari fenomeni biologici si organizzano come i medici avevano previsto, altre volte, invece, questo non succede. La simulazione numerica mette in luce, in questi casi, altre forme di risposte collettive degli attori in gioco, nuovi scenari, inizialmente non previsti. Può anche spiegare risultati apparentemente paradossali. I primi risultati in questa direzione si stanno cominciando a sviluppare in oncologia, neurologia e cardiologia.

Un risultato inaspettato ... ma matematicamente prevedibile

All'inizio di un ictus, i neuroni non hanno abbastanza energia per mantenere le concentrazioni dei vari ioni nei loro valori fisiologici. Molti produttori hanno quindi cercato di sviluppare nuovi farmaci per contrastare questi fenomeni ionici. L'idea era quella di bloccare il passaggio di certi ioni attraverso la parete neuronale, bloccando i canali che questi imboccano. Sono state sviluppate e testate con successo molteplici molecole nei roditori. Ma passando alla sperimentazione sull'uomo, nel 2000, la delusione fu grande: decine di molecole, che rappresentavano anni di lavoro e centinaia di milioni di euro, si sono dimostrate del tutto inefficaci o addirittura dannose! Alcuni test clinici sono stati addirittura interrotti perché inducevano un aumento del tasso di mortalità nell'uomo: la molecola aveva esattamente l'effetto opposto di quello previsto.

Questo fallimento totale si è verificato nonostante i canali ionici siano stati studiati e siano ben noti fin dagli anni cinquanta (Hodgkin e Huxley hanno ricevuto il Premio Nobel per la medicina per questa scoperta). Era possibile prevedere una tale sconfitta? Era possibile prevedere alcune delle difficoltà incontrate attraverso un'analisi matematica?

Il modello matematico di scambio degli ioni mostra che il rapporto cellule gliali / neuroni ha un effetto determinante sulla efficacia dei bloccanti dei canali ionici. Tuttavia, l'effetto di questo rapporto è difficile da prevedere per pura intuizione.

Qui, per costruire il modello matematico, l'approccio è quello di fare prima una lista degli ioni importanti, dei canali ionici, delle pompe, degli scambiatori, ... nei quali questi ioni circolano. Di seguito bisogna stabilire i parametri fisiologici relativi agli oggetti elencati (concentrazioni, numero di porte, potenziale d'apertura ...). Successivamente tutto ciò viene tradotto in equazioni (utilizzando le equazioni di Nernst, i potenziali di membrana ...). Questo lavoro lungo e delicato si basa su un approccio adottato negli ultimi decenni, in particolare per il cuore, sviluppato, tra gli altri, dal team del professor Denis Noble di Oxford.

Il modello costruito indica chiaramente gli effetti positivi dei diversi farmaci nei roditori ... ma anche gli effetti molto deboli o addirittura quelli negativi sull'uomo! Cioè, se negli anni '90 avessimo tradotto in equazioni gli scambi ionici in un ictus ed effettuato delle simulazioni con questi modelli, avremmo potuto prevedere quello che poi è successo.

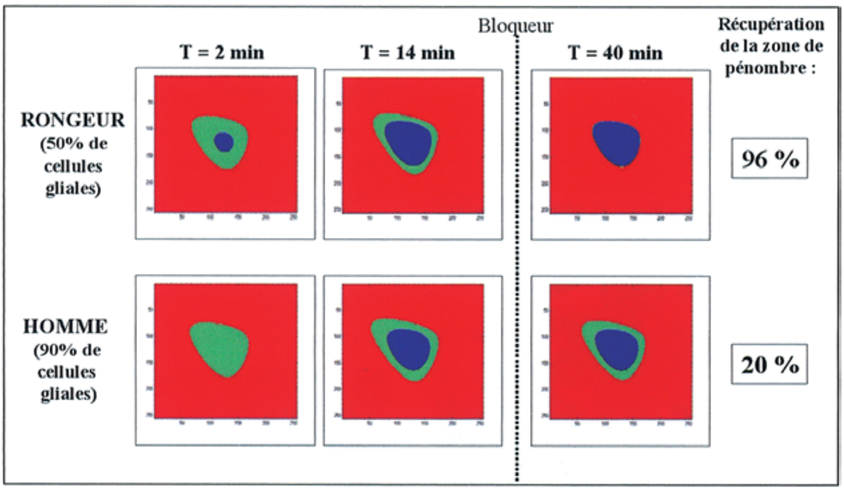
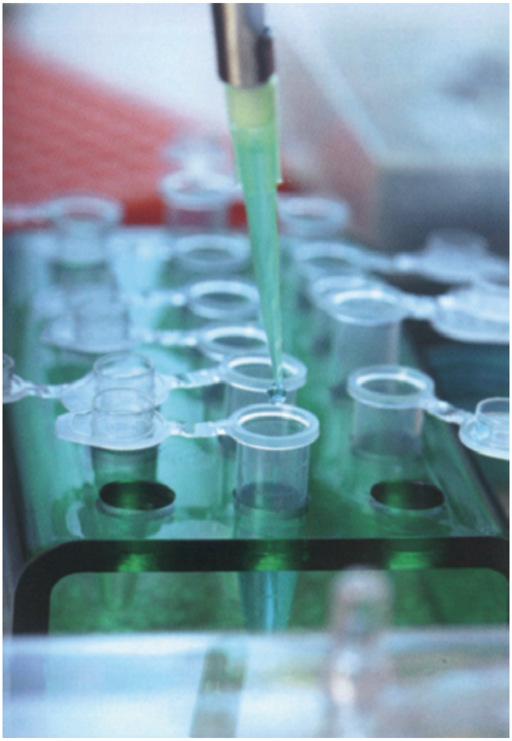


Figura 1: I risultati della simulazione ottenuti con il modello dell'ictus che mostrano l'evoluzione delle diverse aree ischemiche nei 40 minuti successivi all'ostruzione arteriosa (t = 0 min) e dopo l'introduzione di uno specifico bloccante di un canale ionico sodico specifico (t = 20 min) nel "modello di roditore" e nel "modello umano" (diversi per la loro proporzione di cellule gliali). La zona infartuata (dove le cellule sono completamente morte) è in blu, la zona d'ombra (dove le cellule sono danneggiate o indebolite, e possono sia morire, sia riprendersi e sopravvivere) in verde e la zona sana in rosso. L'obiettivo di qualsiasi trattamento è ovviamente quello di fare in modo che la zona di d'ombra si evolva il più possibile verso lo stato sano. I risultati di queste simulazioni mostrano che, a seguito dell'introduzione del bloccante, il recupero della zona d'ombra è considerevole nel caso dei roditori, ma piuttosto debole nel caso dell'uomo. (Tratto da M.A. Dronne, E. Grenier, G. Chapuisat, M. Hommel, J.-P. Boissel, Progress in Biophysics and Molecular Biology, 2008, vol. 97).

Possiamo capire il fallimento clinico a partire da questo modello? Qui, il punto cruciale è la diversa composizione del tessuto nervoso del cervello nell'uomo e nei roditori. Nel cervello, sono presenti neuroni e cellule gliali (cellule di supporto che nutrono i neuroni, li proteggono, contribuiscono alla omeostasi ...). Nei ratti, ci sono 2 cellule gliali per ogni neurone, mentre nell'uomo ci sono circa 9 cellule gliali per ogni neurone. Il modello matematico di scambio ionico mostra che questo rapporto cellule gliali / neuroni ha un effetto determinante sull'efficienza dei bloccanti dei canali ionici: più questo rapporto aumenta e meno il farmaco è efficace. Essendo le cellule gliali proporzionalmente più numerose negli esseri umani, l'effetto del farmaco svanisce e diventa debole o addirittura negativo in alcuni casi. Tuttavia, l'effetto di questa relazione cellule gliali / neuroni è difficile da trovare con la pura intuizione. Per questo motivo l'uso di un modello matematico e la sua simulazione erano necessari per testare l'influenza di questo parametro in modo più preciso e individuare i pericoli di questo scenario inaspettato.



FAQ (e preconcetti) circa i matematici

Quali sono le opportunità per chi studia matematica? Come si diventa un matematico? In che cosa consiste questo lavoro? Professionisti, uomini e donne, rispondono a queste domande e abbattono i preconcetti.

Quale professione si può esercitare con la laurea in matematica?

Edwige Godlewski,
responsabile della laurea in Ingegneria Matematica presso Université Pierre et Marie Curie:

Le aree dove i giovani formati in matematica trovano un impiego coprono quasi tutti i settori: automobilistico, aeronautico, spaziale, energetico, dei trasporti, delle telecomunicazioni, elaborazione dei segnali, trattamento delle immagini, farmaceutico, biomedicale, dell'ingegneria civile e ambientale, logistico, bancario, assicurativo, delle previsioni, dei media, ecc. Nel profilo professionale appare spesso il termine consulente scientifico (se a livello di laurea magistrale) o analista, responsabile di progetto, ecc.[1] Il tipo di incarico è anch'esso molto vario, potendo richiedere in maniera predominante delle competenze matematiche (ad esempio, per professioni che richiedono competenze tecniche molto specialistiche di statistica o di metodi numerici) o limitandosi, talvolta sin dall'inizio, a qualità trasversali, tradizionalmente pertinenti alla disciplina: astrazione, rigore, spirito analitico, ecc.

In generale, la domanda di giovani preparati nelle scienze statistiche è molto significativa e le applicazioni sono piuttosto note (scienze attuariali, biostatistica, statistiche industriali, studi di marketing, previsione dei consumi, sondaggi, epidemiologia, ecc.). Ma ci sono anche reali opportunità per chi segue un master in qualsiasi settore della matematica applicata, a

[1] Qui e nel seguito si è cercato di tradurre i titoli di studio e le posizioni professionali francesi con quelle italiane approssimativamente equivalenti. N.d.T.

condizione che la formazione comprenda un minimo di padronanza degli strumenti informatici e software e che il giovane laureato impari a capire i problemi dal punto di vista dell'impresa, senza rimanere bloccato nell'idea di non poter applicare altro che teoremi. Uno stage in azienda, al termine del corso di laurea, è un passo quasi indispensabile. A queste condizioni, i responsabili della formazione dei master, come quelli in matematica, o che comunque richiedono una laurea magistrale in matematica, constatano che i loro studenti, se mostrano un certo dinamismo, trovano abbastanza rapidamente un lavoro interessante e che la formazione permette loro sia di rimanere in un dominio scientifico che di passare a ruoli di responsabile di progetto o altri tipi di posizioni.

È necessario avere un dottorato per svolgere un lavoro che utilizza la matematica?

Adeline Samson,
responsabile del corso di laurea "Statistique et Informatique décisionelle pour la Santé" presso IUT Paris Descartes:

No! È la ricchezza delle opportunità date dalla matematica. I mestieri che coinvolgono i matematici sono accessibili con tutti i livelli di formazione. La differenza con le professioni accessibili dopo un dottorato di ricerca riguarda il modo in cui viene utilizzata la matematica. Chi ha un dottorato in matematica sarà chiamato a sviluppare o inventare nuovi modelli matematici, nuovi metodi di analisi, di calcolo, di stima statistica. Con la laurea, invece, verrà richiesto di utilizzare la matematica attraverso dei software o dei metodi già sviluppati da ricercatori e soprattutto di saper scegliere il metodo o il modello che meglio si adatta al problema sul quale si lavora.

Facoltà di Jussieu

In che cosa consiste la professione di ricercatore in matematica? Dove lavorano i ricercatori? Che cosa fanno?

Cédric Villani,
Professore presso Université Claude Bernard Lyon 1 e direttore presso Institut Henri Poincaré (CNRS / UPMC):

Nelle università, negli istituti di ricerca e nei centri di collaborazioni scientifiche, i ricercatori in matematica lavorano per sviluppare delle teorie matematiche e per trasmettere le proprie conoscenze. È in questo modo che nuovi teoremi continuano a nascere, ora più che mai. Il numero stimato di nuovi teoremi è ogni anno tra 100.000 e un milione! Questa nascita non è un processo semplice: è il frutto di tentativi ed errori, incontri e discussioni, di lavoro paziente e di idee fulminanti, di ispirazioni ricevute non si sa bene da dove. Gli strumenti del mestiere:

certamente carta e matita; computer, sia per effettuare dei calcoli, che per comunicare; libri, quelli sono sempre indispensabili; scambi, scambi e ancora scambi. I ricercatori sono profondamente consapevoli di formare una comunità pienamente internazionale. Viaggiano continuamente in tutto il mondo per tenere conferenze nelle quali parlano delle loro scoperte, dei loro problemi e delle loro speranze.

Esistono ancora cose da scoprire in matematica?

Filippo Santambrogio,
docente presso Université Paris-Sud:

Naturalmente! E le domande che sorgono sono di natura molto varia. Da una parte ci sono problemi molto difficili, rimasti irrisolti da secoli. Alcuni sono anche molto semplici da raccontare: per esempio, la congettura di Goldbach (dimostrare che tutti i numeri pari maggiori di 2 si possono scrivere come somma di due numeri primi. Ciò è vero per tutti i numeri che i computer sono riusciti a testare, ma nessuno è ancora riuscito a dimostrare che sia sempre vero!).

Dall'altra parte, ci sono problemi che derivano dalle applicazioni. Quando ho iniziato, insieme al mio collega Bertrand Maury, a studiare i modelli di movimento delle folle, abbiamo visto che era necessario considerare il limite di una relazione non lineare tra pressione e densità: tutto questo è molto meno facile da raccontare rispetto alla congettura di Goldbach e non lo avremmo mai considerato interessante prima di vederne le applicazioni... ma è così che la maggior parte la matematica viene sviluppata: studiamo quello di cui qualcuno ha bisogno, senza saperlo in anticipo!

Bisogna essere un genio per diventare matematico?

Secondo Terence Tao, uno dei migliori matematici contemporanei (vincitore della medaglia Fields nel 2006): "La risposta è no, assolutamente no. Portare dei contributi belli e utili alla matematica richiede molto lavoro, specializzarsi in un settore, imparare cose in altri settori, porre domande, parlare con altri matematici e riflettere sui grandi orizzonti del paesaggio matematico considerato. E, sì, naturalmente, un'intelligenza ragionevole, pazienza e maturità sono ovviamente necessarie. Ma, in nessun caso, è necessario possedere una sorta di gene magico del genio matematico o altri superpoteri, che forniscano ispirazioni spontanee o che facciano sorgere dal nulla delle idee profonde o delle soluzioni totalmente inaspettate ai problemi."

Peraltro, solo in Francia, circa 4.000 matematici lavorano nei laboratori di ricerca delle università, per non parlare di tutti i matematici che lavorano nelle aziende... certamente non tutti dei geni!

Che cosa è un dottorato di ricerca in matematica?

Amélie Rambaud,
dottorato presso l'Institut Camille Jordan
sotto la direzione di Francis Filbet:

Tre anni, che sembrano lunghi, ma è quello che ci vuole! Per esplorare, prendere confidenza con il mondo della ricerca (in matematica, intendo!). Si tratta di affrontare un problema di matematica aperto proposto da un relatore, ma questo può portare, strada facendo, ad altre piste, ad esempio nel caso in cui l'argomento iniziale sollevi questioni che paiono più rilevanti. In realtà tutto ciò è piuttosto soggettivo, istintivo. In una tesi di dottorato, ci sono spesso momenti in cui ci si trova persi, per esempio impelagati nei calcoli. Questa nebbia può durare a lungo! Ma perseverando, lavorando, arriva un momento in cui tutto si sblocca e avviene la "nascita" di un risultato, questa almeno è la mia impressione! Poi arriva una fase di redazione, l'esposizione dei risultati ad altri ricercatori. Questo è un esercizio difficile, che deve essere svolto in inglese, ma molto interessante, che permette di collocare il proprio lavoro in un contesto, metterlo in prospettiva, insomma, trovare il suo posto nella proliferazione di pubblicazioni matematiche. Alla fine del percorso si ottiene una grande soddisfazione: vedere la propria opera pubblicata è una conquista che dà la motivazione per continuare la propria ricerca! Direi, quindi, che fare una tesi di dottorato significa mettere un piede nel pazzo mondo della ricerca. Qualunque sia il risultato, è un'esperienza che arricchisce, ma, attenzione: si rischia di rimanere infettati dal virus e di continuare con la ricerca!

È possibile lavorare nell'industria, con una laurea in matematica?

Marc Bernot,
ricercatore presso Thales Alenia Space, Cannes:

Certamente! Mi sono unito alla società di ricerca spaziale Thales Alenia (società che produce satelliti di vario tipo) da quattro anni. Lavoro sulla qualità ottica dei telescopi, settore che mi mette in contatto con colleghi provenienti da ambienti diversi (ottica, meccanica, ...). Questa diversità è stata molto stimolante dopo la mia formazione al 100% in matematica (Scuola Normale Superiore, abilitazione all'insegnamento, master, dottorato), senza passare da una scuola di ingegneria.

Ma che cosa ci fanno i matematici con i telescopi, se la loro modellizzazione si basa sulla fisica? Il fatto è che, spesso, la conoscenza del modello da sola non basta ... Un passo importante è la modellizzazione numerica e la sua messa in opera in modo rapido, preciso e coinvolge una serie di problematiche. Le tecniche non sono sempre quelle assai complesse che ho visto durante la mia formazione: è il contesto applicativo che è complesso e interdisciplinare. Eppure la mia formazione (persino la teoria di Galois, della quale non mi servo assolutamente più) è essenziale: è grazie a lei che non provo timore di fronte a nessun tecnicismo, ma un desiderio di chiarire e trovare ciò che è essenziale.

Sébastien Marque,
direttore del Dipartimento Biometria di Danone Research:

Sì! Per quello che mi riguarda ho una laurea in Matematica Applicata e un dottorato di ricerca in Biostatistica. Delle venti persone che lavorano nel mio reparto, due terzi hanno una laurea nel campo della statistica o della biostatistica. La maggior parte delle altre persone sono data manager: professione che si può esercitare a diversi livelli e con diversi gradi.

Oltre alle competenze tecniche in matematica/statistica ottenute con la laurea, mi aspetto dalle persone che lavorano nella mia squadra una qualche apertura mentale. Sono queste attitudini che fanno da ponte per tradurre una questione scientifica (biologica o epidemiologia molto spesso) in un problema statistico e, una volta risolto, devono di nuovo fare da ponte per comunicare le risposte ottenute in modo semplice, affinché i risultati siano comprensibili e utilizzabili. Inoltre, il mondo sta cambiando rapidamente, devono, quindi, costantemente rinnovarsi e adattarsi, sia per quello che riguarda l'organizzazione aziendale, sia nel campo della matematica e della statistica.

Mi piace la matematica e mi piace molto lavorare in gruppo, è compatibile?

Lavorare da soli o in gruppo: questo dipende principalmente dalla personalità di ogni ricercatore. Vengono organizzati continuamente incontri nel campo della ricerca e la fermentazione e lo scambio di idee sono cruciali per lo sviluppo della ricerca matematica. Le migliori idee possono provenire sia da una vivace discussione dopo il pasto, come dall'isolamento di una notte insonne. Il lavoro, anche se svolto singolarmente, è in qualche modo non isolato, non fosse altro per le comunicazioni elettroniche e per i progressi scientifici che si costruiscono uno sull'altro ...

La ricerca matematica è compatibile con la famiglia e i figli?

Fabienne Castell,
professoressa universitaria a Aix-Marseille,
tre figli:

Deve essere compatibile, perché, di fatto, io sono sia una madre che un matematico, ma è difficile per me giudicare dall'interno la qualità della mia vita familiare! Quello che posso dire è che non sento che la matematica "sacrifichi" la mia vita familiare e, in ogni caso, né più né meno di tanti altri lavori che avrei potuto fare. Che tu sia uomo o donna, c'è sempre un equilibrio tra vita familiare e vita professionale e questo problema non è specifico per il matematico (uomo o donna). Per esempio, non ho calcolato o organizzato la mia vita (mi riferisco essenzialmente alla nascita dei miei figli) rispetto alle tappe più importanti della mia professione. Io, sicuramente, mi muovo di meno rispetto alle persone che non hanno famiglia e se, certamente, questo ha un impatto sulla mia carriera, non è poi così importante. Il mio lavoro è impegnativo e miei figli lo sanno e a ciò sono abituati. Ciò che è specifico di questa professione è che si tratta di un lavoro creativo e, talvolta, anche a casa, la mia mente è occupata dalla mia ricerca e faccio fatica a pensare ad altro.

Non riesco a decidermi tra matematica e biologia: posso studiare biologia all'interno di un curriculum in matematica?

Hélène Morlon
ricercatrice in biologia presso il Centre de Mathématique Appliquées de l'École polytechnique:

Sì, certamente. Sempre più le aree della biologia si basano su modelli e vengono ricercate persone con forti competenze matematiche e interesse per la biologia. Un buon approccio può essere quello di promuovere, in parallelo con la matematica pura o applicata, la formazione in statistica, analisi numerica, programmazione ... e naturalmente biologia! Appassionata di biologia, ma anche molto amante della matematica, ho scelto un corso di matematica, dopo la scuola superiore, per lasciarmi la massima possibilità di scelta. Dopo una laurea di primo livello in matematica ho studiato presso l'École Normale Supérieure di Cachan (dove ho conseguito anche abilitazione all'insegnamento in matematica), ho potuto integrare un master in ecologia e iniziare una conversione alla biologia. È sicuramente una sfida, ma l'interdisciplinarità è sempre più apprezzata. Ne vale la pena, in quanto permette di combinare diverse passioni! Molti dei miei colleghi hanno seguito un percorso simile e non credo che abbiano alcun rimpianto.

Un matematico può creare la propria impresa?

Stéphane Mallat,
professore presso la Scuola Normale
e creatore della società Let it Wave:

Chiaramente sì, e molti matematici l'hanno fatto con successo sia Francia che all'estero, in particolare negli Stati Uniti. La matematica è una straordinaria fonte di innovazione tecnologica per sviluppare nuovi prodotti e migliorare i processi industriali. Ciò avviene in informatica, nelle telecomunicazioni, nell'ambito aerospaziale, in quello sanitario e nella maggior parte dei settori industriali.

La professione di docente-ricercatore è un'eccellente formazione per diventare un imprenditore, contrariamente a molti pregiudizi. Avviare un'impresa è anche un progetto di ricerca. Partiamo da una prima idea, per poi affrontare la ricerca di mercato, l'implementazione della tecnologia, la creazione di una squadra con molte sorprese e cambiamenti di rotta. L'esperienza di insegnamento aiuta a spiegare e convincere del valore delle proprie idee.

Non è necessario conoscere la finanza, il commercio, il marketing o il management, ma semplicemente, aver voglia di imparare, di scoprire nuovi orizzonti. Ecco quello che ho vissuto.

Come ci si guadagna da vivere con la professione del matematico?

Ovviamente, gran parte dipende dal dominio di competenza ... Gli stipendi di partenza dei giovani matematici laureati che si orientano verso la finanza sono certamente interessanti, anche se un elemento di mistero (o di fantasia) viene mantenuto riguardo al livello reale degli stipendi in questa area. In confronto, le carriere nel settore pubblico (insegnamento universitario, ricerca in istituti pubblici ...) sono ovviamente meno redditizie. Da un estremo all'altro, i guadagni sembrano in definitiva molto vari, dall'insegnamento alla banca passando per la ricerca industriale o ambientale e della salute ...